D2

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :
C12N 15/82, C07K 14/415, C12N 5/10, 15/62, A01H 5/00

A2

(11) International Publication Number: WO 00/12732

(43) International Publication Date: 9 March 2000 (09.03.00)

(21) International Application Number: PCT/US99/18955

(22) International Filing Date: 25 August 1999 (25.08.99)

(30) Priority Data:
60/098,225 28 August 1998 (28.08.98) US

(71) Applicant *(for all designated States except US)*: PIONEER HI–BRED INTERNATIONAL, INC. [US/US]; 800 Capital Square, 400 Locust Street, Des Moines, IA 50319 (US).

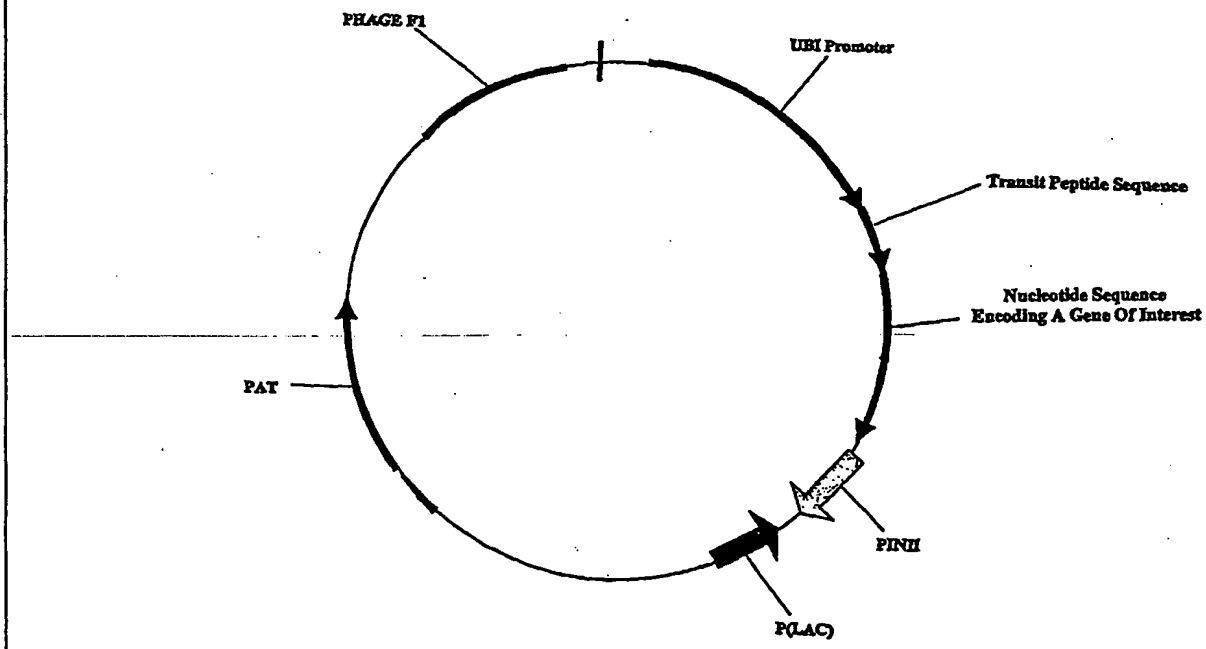
(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: BENSEN, Robert, J. [US/US]; 200 Northwest Prairie Creek Drive, Grimes, IA 50111 (US).

(74) Agents: SPRUILL, W., Murray et al.; Alston & Bird LLP, P.O. Drawer 34009, Charlotte, NC 28234–4009 (US).

(81) Designated States: AE, AL, AM, AT, AT (Utility model), AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, CZ (Utility model), DE, DE (Utility model), DK, DK (Utility model), DM, EE, EE (Utility model), ES, FI, FI (Utility model), GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SK (Utility model), SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: ORGANELLE TARGETING SEQUENCES



(57) Abstract

Compositions and methods are provided for modulating the subcellular localization of proteins in a cell. Compositions include nucleotide and amino acid sequences of transit peptide sequences from maize. Such sequences find utility in the enhanced or modified localization of protein to a plastid or compartment thereof.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania
AM	Armenia
AT	Austria
AU	Australia
AZ	Azerbaijan
BA	Bosnia and Herzegovina
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CU	Cuba
CZ	Czech Republic
DE	Germany
DK	Denmark
EE	Estonia
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GH	Ghana
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IL	Israel
IS	Iceland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgyzstan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakstan
LC	Saint Lucia
LI	Liechtenstein
LK	Sri Lanka
LR	Liberia
LS	Lesotho
LT	Lithuania
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
MK	The former Yugoslav Republic of Macedonia
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
MX	Mexico
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SG	Singapore
SI	Slovenia
SK	Slovakia
SN	Senegal
SZ	Swaziland
TD	Chad
TG	Togo
TJ	Tajikistan
TM	Turkmenistan
TR	Turkey
TT	Trinidad and Tobago
UA	Ukraine
UG	Uganda
US	United States of America
UZ	Uzbekistan
VN	Viet Nam
YU	Yugoslavia
ZW	Zimbabwe

ORGANELLE TARGETING SEQUENCES

FIELD OF THE INVENTION

The invention is drawn to the genetic modification of plants, particularly to the targeting of proteins to cellular organelles.

BACKGROUND OF THE INVENTION

Plastids are a class of plant organelles derived from proplastids and include chloroplasts, leucoplasts, amyloplasts, and chromoplasts. The plastids are major sites of biosynthesis in plants. In addition to photosynthesis in the chloroplast, plastids are also sites of lipid biosynthesis, nitrate reduction to ammonium, and starch storage. And while plastids contain their own circular genome, most of the proteins localized to the plastids are encoded by the nuclear genome and are imported into the organelle from the cytoplasm.

The mechanism of protein import into the plastids has been most extensively studied in the chloroplast. The chloroplast is a complex cellular organelle composed of three membranes: the inner envelope membrane, the outer envelope membrane, and the thylakoid membrane. The membranes together enclose three aqueous compartments termed the intermediate space, the stroma, and the thylakoid lumem.

Proteins imported from the cytosol generally contain, at their amino terminus, short sequences referred to as "transit peptides" that are responsible for post-translational targeting of the protein to the chloroplast. The import process is initiated by binding of precursor proteins to the chloroplast surface, followed by the subsequent translocation of the precursor protein across the chloroplast envelope membranes. The transit peptide is typically an expendable part of the protein, and upon translocation into the chloroplast the amino acid sequence is cleaved from the precursor protein. Further sub-organellar sorting of the modified precursor takes place as appropriate.

Genes reported to have naturally encoded transit peptide sequences at their N-terminus include the chloroplast small subunit of ribulose-1,5-bisphosphate carboxylase (Rubisco),de Castro Silva Filho *et al.* (1996) *Plant Mol. Biol. 30*:769-

780; Schnell, D.J. *et al.* (1991) *J. Biol. Chem. 266(5)*:3335-3342; 5-(enolpyruvyl)shikimate-3-phosphate synthase (EPSPS), Archer *et al.* (1990) *J. Bioenerg. and Biomemb. 22(6)*:789-810; tryptophan synthase, Zhao, J. *et al.* (1995) *J. Biol. Chem. 270(11)*:6081-6087, plastocyanin, Lawrence *et al.* (1997) *J. Biol. Chem. 272(33)*:20357-20363, chorismate synthase, Schmidt *et al.* (1993) *J. Biol. Chem. 268(36)*:27477-27457, and the light harvesting chlorophyll a/b binding protein(LHBP), Lamppa *et al.* (1988) *J. Biol. Chem. 263*:14996-14999.

Statistical analysis of transit peptides that direct protein localization to chloroplasts has revealed a sequence profile for transit peptides with the following characteristics. In the central region, the peptides typically contain an exceptionally high content of basic and hydroxylated amino acids, such as serine and threonine. In addition, there is a near absence of negatively charged amino acids such as aspartic acid, glutamic acid, asparagine, and glutamine. The amino-terminal region is devoid of charged amino acids and lacks turn promoting amino acids such as glycine and proline. The carboxy terminal domain is high in arginine and has a capacity for forming an amphipathic beta sheet secondary structure. The length of the transit peptide is variable, commonly between 50 and 120 amino acids. In addition, there is a well conserved cleavage site (V/I)X(C/A)$^{\downarrow}$A. Often one or more arginines are found some 5 to 10 residues upstream of this cleavage site. Exceptions to the transit peptide signals described above are known. See von Heijne *et al.* (1989) *Eur. J. Biochem 180*:535-545.

Because proteins containing transit peptides are localized to the chloroplast with a high degreed of specificity (Boutry *et al.* (1987) *Nature 328*:340-342; de Boer *et al.* (1991) *Biochem. Biophys. Acta 1071*:221-253), transit peptide sequences prove useful in recombinant DNA technology. For example, transit peptide sequences may be inserted into an expression cassette and serve to guide the expressed protein to the chloropast. In plants, transit peptide signals have been useful in the localization of proteins responsible for herbicide or antibacterial resistance to the chloroplast.

Although transit peptides have been described, only a few have been utilized successfully in attempts to target chimeric molecules to chloroplasts. Thus, there is a need for additional DNA sequences that encode transit peptides for

use in future genetic engineering projects that require specific targeting of foreign proteins to chloroplast.

SUMMARY OF THE INVENTION

The present invention provides methods and compositions for the subcellular localization of proteins. Specifically, the invention provides a means to direct the localization of a protein to a plant cell organelle, more particularly to a plant plastid. Compositions of the present invention include the nucleotide and amino acid sequences of novel plastid targeting sequences, hence referred to as transit peptides. Such sequences find utility in the enhanced or modified localization of proteins to a plastid or compartment thereof.

Further compositions of the invention include, expression cassettes and transformation vectors comprising the isolated nucleotide sequences of the transit peptides. Also provided are transgenic plants, plant cells, and plant tissue that express proteins that have been localized to a plastid using the transit peptides of the present invention.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 schematically illustrates a plasmid vector comprising a ubiquitin promoter operably linked to a transit peptide sequence of the invention operably linked to a gene of interest.

DETAILED DESCRIPTION OF THE INVENTION

Compositions and methods are provided for modulating the subcellular localization of proteins. Specifically, the compositions of the invention include maize transit peptide sequences that find use in modulating the cellular localization of a protein of interest. In particular, the transit peptides of the invention finds use in the localization of proteins to plant organelles, particularly to plastids and compartments thereof.

By "plastid" is intended a class of plant cell organelles comprising proplastids, leucoplasts, amyloplast, chromoplasts, and chloroplast. By "plastid or compartment thereof" is intended any plastid structure, membrane or compartment of a plastid. For example, when referring to a chloroplast, "a

compartment thereof" encompasses the intermediate envelope space, the stroma, the lumen, the outer envelope, the inner envelope and the thylakoid membrane.

A signal or targeting sequence is a structural peptide domain required for targeting of a given polypeptide to a subcellular organelle, subcellular compartment or secretion from the cell. The transport of a protein of interest to a subcellular compartment is accomplished by operably linking the nucleotide sequence encoding a signal sequence to the 5' and/or 3' region of the gene encoding the protein of interest. During protein synthesis and processing, the targeting sequence influence where the protein of interest is ultimately compartmentalized.

By "transit peptide" is intended a polypeptide that directs the transport of a nuclear encoded protein to a plastid or a compartment thereof. Typically, the transit peptide sequence is located at the amino-terminus of a polypeptide. However, the transit peptide may also be located at either the c-terminus or internally in the polypeptide.

The maize sequences provided by the present invention includes a maize transit peptide having homology to the maize light harvesting chlorophyll a/b binding protein (SEQ ID NOS: 1 and 2). The present invention also provides a transit peptide having a homology to the maize ribulose bisphosphate carboxylase/oxygenase protein (SEQ ID NOS: 3 and 4).

Also provided are maize transit peptide sequences that share homology to transit peptide sequences of various non-maize gene products including, EPSP synthase (SEQ ID NOS: 5 and 6), tryptophan synthase component (SEQ ID NOS: 7 and 8), ribosomal protein L35 (SEQ ID NOS: 9 and 10), plastid ribosomal protein CL9 (SEQ ID NOS: 11 and 12), plastocyanin (SEQ ID NOS: 13 and 14), 3-dehydroquinate synthase (SEQ ID NOS: 15 and 16), plastid ribosomal protein CL15 (SEQ ID NOS: 17 and 18), chorismate synthase (SEQ ID NOS: 19 and 20), and choporphyringogen oxidase (SEQ ID NOS: 21 and 22).

In particular, the present invention provides for isolated nucleic acid molecules comprising nucleotide sequences encoding the amino acid sequences shown in SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, and 22. Further provided are polypeptides having an amino acid sequence encoded by a nucleic acid

molecule described herein, for example those set forth in SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, and 21 and fragments and variants thereof.

The invention encompasses isolated or substantially purified nucleic acid or polypeptide compositions. An "isolated" or "purified" nucleic acid molecule or protein, or biologically active portion thereof, is substantially free of other cellular material, or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized. Preferably, an "isolated" nucleic acid is free of sequences (preferably protein encoding sequences) that naturally flank the nucleic acid (i.e., sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated nucleic acid molecule can contain less than about 5 kb, 4 kb, 3 kb. 2 kb, 1 kb, 0.5 kb, or 0.1 kb of nucleotide sequences that naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. A polypeptide that is substantially free of cellular material includes preparations of polypeptides having less than about 30%, 20%, 10%, 5%, (by dry weight) of contaminating polypeptides. When the polypeptide of the invention or biologically active portion thereof is recombinantly produced, preferably culture medium represents less than about 30%, 20%, 10%, or 5% (by dry weight) of chemical precursors or non-protein-of-interest chemicals.

Fragments and variants of the disclosed nucleotide sequences and the polypeptides encoded thereby are also encompassed by the present invention. By "fragment" is intended a portion of the nucleotide sequence or a portion of the amino acid sequence and hence polypeptide encoded thereby. Fragments of a nucleotide sequence may encode polypeptide fragments that retain the biological activity of the native polypeptide and hence facilitates the transport of a nuclear encoded protein to a plastid or a compartment thereof. Alternatively, fragments of a nucleotide sequence that are useful as hybridization probes generally do not encode fragment polypeptides retaining biological activity. Thus, fragments of a nucleotide sequence may range from at least about 20 nucleotides, about 50 nucleotides, about 100 nucleotides, and up to the full-length nucleotide sequence encoding the proteins of the invention.

A fragment of a transit peptide nucleotide sequence that encodes a biologically active portion of a transit peptide of the invention will encode at least 15, 25, 30, 40, 50, 60, 70, 80, 90, 100 contiguous amino acids, or up to the total number of amino acids present in a full-length transit peptide of the invention (for example, 135, 117, 144, 152, 150, 152, 145, 132, 134, 166, and 107 amino acids for SEQ ID NOS; 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, and 22, respectfully). Fragments of a transit peptide nucleotide sequence that are useful as hybridization probes for PCR primers generally need not encode a biologically active portion of a transit peptide.

Thus, a fragment of a transit peptide nucleotide sequence may encode a biologically active portion of a transit peptide, or it may be a fragment that can be used as a hybridization probe or PCR primer using methods disclosed below. A biologically active portion of a transit peptide can be prepared by isolating a portion of one of the transit peptide nucleotide sequences of the invention, expressing the encoded portion of the transit peptide(e.g., by recombinant expression *in vitro*), and assessing the activity of the encoded portion of the transit peptide. Nucleic acid molecules that are fragments of a transit peptide nucleotide sequence comprise at least 16, 20, 50, 75, 100, 150, 200, 250, 300, 350, 400 nucleotides, or up to the number of nucleotides present in a full-length transit peptide nucleotide sequence disclosed herein (for example, 407, 426, 455, 459, 560, 461, 437, 463, 448, 528, and 427 nucleotides for SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, and 21, respectively).

By "variants" is intended substantially similar sequences. For nucleotide sequences, conservative variants include those sequences that, because of the degeneracy of the genetic code, encode the amino acid sequence of one of the transit peptide of the invention. Naturally occurring variants such as these can be identified with the use of well-known molecular biology techniques, as, for example, with polymerase chain reaction (PCR) and hybridization techniques as outlined below. Variant nucleotide sequences also include synthetically derived nucleotide sequences, such as those generated, for example, by using site-directed mutagenesis but which still encode a transit peptide or protein of the invention. Generally, nucleotide sequence variants of the invention will have at least 40%,

50%, 60%, 70%, generally, 80%, preferably 85%, 90%, up to 95%, 98% sequence identity to its respective native nucleotide sequence.

By "variant" polypeptide is intended as a polypeptide derived from the native polypeptide by deletion (so-called truncation) or addition of one or more amino acids to the N-terminal and/or C-terminal end of the native polypeptide; deletion or addition of one or more amino acids at one or more sites in the native polypeptide; or substitution of one or more amino acids at one or more sites in the native protein. Such variants may result from, for example, genetic polymorphism or from human manipulation.

The transit polypeptides of the invention may be altered in various ways including amino acid substitutions, deletions, truncations, and insertions. Methods for such manipulations are generally known in the art. For example, amino acid sequence variants of the transit peptide can be prepared by mutations in the DNA. Methods for mutagenesis and nucleotide sequence alterations are well known in the art. See, for example, Kunkel (1985) *Proc. Natl. Acad. Sci. USA 82*:488-492; Kunkel *et al.* (1987) *Methods in Enzymol. 154*:367-382; U.S. Patent No. 4,873,192; Walker and Gaastra, eds. (1983) *Techniques in Molecular Biology* (MacMillan Publishing Company, New York) and the references cited therein. Guidance as to appropriate amino acid substitutions that do not affect biological activity of the protein of interest may be found in the model of Dayhoff *et al.* (1978) *Atlas of Protein Sequence and Structure* (Natl. Biomed. Res. Found., Washington, D.C.), herein incorporated by reference. Conservative substitutions, such as exchanging one amino acid with another having similar properties, may be preferred.

Thus, the nucleotide sequences of the invention include both the naturally occurring sequences as well as mutant forms. Likewise, the polypeptides of the invention encompass both naturally occurring polypeptides as well as variations and modified forms thereof. Such variants will continue to possess the desired ability to facility the transport of a nuclear encoded protein to a plastid or compartment thereof. Obviously, the mutations that will be made in the DNA encoding the variant must not place the sequence out of reading frame and preferably will not create complementary regions that could produce secondary mRNA structure. See, EP Patent Application Publication No. 75,444.

The deletions, insertions, and substitutions of the polypeptides sequences encompassed herein are not expected to produce radical changes in the characteristics of the polypeptide. However, when it is difficult to predict the exact effect of the substitution, deletion, or insertion in advance of doing so, one skilled in the art will appreciate that the effect will be evaluated by routine screening assays. That is, the activity can be evaluated by the ability of the isolated sequences to target and deliver a reporter protein to a plastid or compartment thereof. See, for example, de Castro Silva Filho *et al.* (1996) *Plant Mol. Biol.* 30: *796*-780, herein incorporated by reference.

Variant nucleotide sequences and polypeptides also encompass sequences and polypeptides derived from a mutagenic and recombinogenic procedure such as DNA shuffling. With such a procedure, one or more different transit peptide sequences can be manipulated to create a new transit peptide sequence possessing the desired properties. In this manner, libraries of recombinant polynucleotides are generated from a population of related sequence polynucleotides comprising sequence regions that have substantial sequence identity and can be homologously recombined *in vitro* or *in vivo*. For example, using this approach, sequence motifs encoding a domain of interest may be shuffled between the transit peptide sequences of the invention and other known signal sequences or transit peptide sequences to obtain a new nucleotide sequence coding for a transit peptide with an improved property of interest, such as an increased K_m or an increased efficiency and/or specificity of plastid targeting. Strategies for such DNA shuffling are known in the art. See, for example, Stemmer (1994) *Proc. Natl. Acad. Sci. USA* *91*:10747-10751; Stemmer (1994) *Nature 370*:389-391; Crameri *et al.* (1997) *Nature Biotech. 15*:436-438; Moore *et al.* (1997) *J. Mol. Biol. 272*:336-347; Zhang *et al. (1997) Proc. Natl. Acad. Sci. USA 94*:4504-4509; Crameri *et al.* (1998) *Nature 391*:288-291; and U.S. Patent Nos. 5,605,793 and 5,837,458.

The nucleotide sequences of the invention can be used to isolate coresponding sequences from other organisms, particularly other plants, more particularly other monocots. In this manner, methods such as PCR, hybridization, and the like can be used to identify such sequences based on their sequence homology to the sequences set forth herein. Sequences isolated based on their

sequence identity to the entire transit peptide sequences set forth herein or to fragments thereof are encompassed by the present invention.

In a PCR approach, oligonucleotide primers can be designed for use in PCR reactions to amplify corresponding DNA sequences from cDNA or genomic DNA extracted from any plant of interest. Methods for designing PCR primers and PCR cloning are generally known in the art and are disclosed in Sambrook *et al.* (1989) Molecular Cloning: A Laboratory Manual (2d ed., Cold Spring Harbor Laboratory Press, Plainview, New York). See also Innis *et al.*, eds. (1990) PCR Protocols: A Guide to Methods and Applications (Academic Press, New York); Innis and Gelfand, eds. (1995) PCR Strategies (Academic Press, New York); and Innis and Gelfand, eds. (1999) PCR Methods Manual (Academic Press, New York). Known methods of PCR include, but are not limited to, methods using paired primers, nested primers, single specific primers, degenerate primers, gene-specific primers, vector-specific primers, partially-mismatched primers, and the like.

In hybridization techniques, all or part of a known nucleotide sequence is used as a probe that selectively hybridizes to other corresponding nucleotide sequences present in a population of cloned genomic DNA fragments or cDNA fragments (i.e., genomic or cDNA libraries) from a chosen organism. The hybridization probes may be genomic DNA fragments, cDNA fragments, RNA fragments, or other oligonucleotides, and may be labeled with a detectable group such as ^{32}P, or any other detectable marker. Thus, for example, probes for hybridization can be made by labeling synthetic oligonucleotides based on the transit peptide sequences of the invention. Methods for preparation of probes for hybridization and for construction of cDNA and genomic libraries are generally known in the art and are disclosed in Sambrook *et al.* (1989) Molecular Cloning: A Laboratory Manual (2d ed., Cold Spring Harbor Laboratory Press, Plainview, New York).

For example, the entire nucleotide sequence encoding a transit peptide disclosed herein, or one or more portions thereof, may be used as a probe capable of specifically hybridizing to corresponding transit peptide sequences and messenger RNAs. To achieve specific hybridization under a variety of conditions, such probes include sequences that are unique among transit peptide sequences and are preferably at least about 10 nucleotides in length, and most preferably at least

about 20 nucleotides in length. Such probes may be used to amplify corresponding transit peptide sequences from a chosen plant by PCR. This technique may be used to isolate additional coding sequences from a desired plant or as a diagnostic assay to determine the presence of coding sequences in a plant. Hybridization techniques include hybridization screening of plated DNA libraries (either plaques or colonies; see, for example, Sambrook *et al.* (1989) Molecular Cloning: A Laboratory Manual (2d ed., Cold Spring Harbor Laboratory Press, Plainview, New York).

Hybridization of such sequences may be carried out under stringent conditions. By "stringent conditions" or "stringent hybridization conditions" is intended conditions under which a probe will hybridize to its target sequence to a detectably greater degree than to other sequences (e.g., at least 2-fold over background). Stringent conditions are sequence-dependent and will be different in different circumstances. By controlling the stringency of the hybridization and/or washing conditions, target sequences that are 100% complementary to the probe can be identified (homologous probing). Alternatively, stringency conditions can be adjusted to allow some mismatching in sequences so that lower degrees of similarity are detected (heterologous probing). Generally, a probe is less than about 1000 nucleotides in length, preferably less than 500 nucleotides in length.

Typically, stringent conditions will be those in which the salt concentration is less than about 1.5 M Na ion, typically about 0.01 to 1.0 M Na ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes (e.g., 10 to 50 nucleotides) and at least about 60°C for long probes (e.g., greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. Exemplary low stringency conditions include hybridization with a buffer solution of 30 to 35% formamide, 1 M NaCl, 1% SDS (sodium dodecyl sulphate) at 37°C, and a wash in 1X to 2X SSC (20X SSC = 3.0 M NaCl/0.3 M trisodium citrate) at 50 to 55°C. Exemplary moderate stringency conditions include hybridization in 40 to 45% formamide, 1.0 M NaCl, 1% SDS at 37°C, and a wash in 0.5X to 1X SSC at 55 to 60°C. Exemplary high stringency conditions include hybridization in 50% formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 0.1X SSC at 60 to 65°C.

Specificity is typically the function of post-hybridization washes, the critical factors being the ionic strength and temperature of the final wash solution. For DNA-DNA hybrids, the T_m can be approximated from the equation of Meinkoth and Wahl (1984) Anal. Biochem. 138:267-284: $T_m = 81.5°C + 16.6$ (log M) + 0.41 (%GC) - 0.61 (% form) - 500/L; where M is the molarity of monovalent cations, %GC is the percentage of guanosine and cytosine nucleotides in the DNA, % form is the percentage of formamide in the hybridization solution, and L is the length of the hybrid in base pairs. The T_m is the temperature (under defined ionic strength and pH) at which 50% of a complementary target sequence hybridizes to a perfectly matched probe. T_m is reduced by about 1°C for each 1% of mismatching; thus, T_m, hybridization, and/or wash conditions can be adjusted to hybridize to sequences of the desired identity. For example, if sequences with $\geq$90% identity are sought, the T_m can be decreased 10°C. Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point (T_m) for the specific sequence and its complement at a defined ionic strength and pH. However, severely stringent conditions can utilize a hybridization and/or wash at 1, 2, 3, or 4°C lower than the thermal melting point (T_m); moderately stringent conditions can utilize a hybridization and/or wash at 6, 7, 8, 9, or 10°C lower than the thermal melting point (T_m); low stringency conditions can utilize a hybridization and/or wash at 11, 12, 13, 14, 15, or 20°C lower than the thermal melting point (T_m). Using the equation, hybridization and wash compositions, and desired T_m, those of ordinary skill will understand that variations in the stringency of hybridization and/or wash solutions are inherently described. If the desired degree of mismatching results in a T_m of less than 45°C (aqueous solution) or 32°C (formamide solution), it is preferred to increase the SSC concentration so that a higher temperature can be used. An extensive guide to the hybridization of nucleic acids is found in Tijssen (1993) Laboratory Techniques in Biochemistry and Molecular Biology-Hybridization with Nucleic Acid Probes, Part I, Chapter 2 (Elsevier, New York); and Ausubel *et al.*, eds. (1995) Current Protocols in Molecular Biology, Chapter 2 (Greene Publishing and Wiley-Interscience, New York). See Sambrook *et al.* (1989) Molecular Cloning: A Laboratory Manual (2d ed., Cold Spring Harbor Laboratory Press, Plainview, New York).

In general, sequences that encode a transit peptide and hybridize to the nucleic acid sequences encoding a transit peptide disclosed herein will be at least 40% to 50% homologous, about 60% to 70% homologous, and even about 80%, 85%, 90%, 95% to 98% homologous or more with the disclosed sequences. That is, the sequence similarity of sequences may range, sharing at least about 40% to 50%, about 60% to 70%, and even about 80%, 85%, 90%, 95% to 98% sequence similarity.

The following terms are used to describe the sequence relationships between two or more nucleic acids or polynucleotides: (a) "reference sequence", (b) "comparison window", (c) "sequence identity", (d) "percentage of sequence identity", and (e) "substantial identity".

(a) As used herein, "reference sequence" is a defined sequence used as a basis for sequence comparison. A reference sequence may be a subset or the entirety of a specified sequence; for example, as a segment of a full-length cDNA or gene sequence, or the complete cDNA or gene sequence.

(b) As used herein, "comparison window" makes reference to a contiguous and specified segment of a polynucleotide sequence, wherein the polynucleotide sequence in the comparison window may comprise additions or deletions (i.e., gaps) compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. Generally, the comparison window is at least 20 contiguous nucleotides in length, and optionally can be 30, 40, 50, 100, or longer. Those of skill in the art understand that to avoid a high similarity to a reference sequence due to inclusion of gaps in the polynucleotide sequence a gap penalty is typically introduced and is subtracted from the number of matches.

Methods of alignment of sequences for comparison are well known in the art. Optimal alignment of sequences for comparison may be conducted by the local homology algorithm of Smith *et al.* (1981) *Adv. Appl. Math. 2*:482; by the homology alignment algorithm of Needleman *et al.* (1970) *J. Mol. Biol. 48*:443; by the search for similarity method of Pearson *et al.* (1988) *Proc. Natl. Acad. Sci. 85*:2444; by computerized implementations of these algorithms, including, but not limited to: CLUSTAL in the PC/Gene program by Intelligenetics, Mountain View, California; GAP, BESTFIT, BLAST, FASTA, and TFASTA in the Wisconsin

Genetics Software Package, Genetics Computer Group (GCG), 575 Science Drive, Madison, Wisconsin, USA; the CLUSTAL program is well described by Higgins *et al.* (1988) *Gene 73*:237-244 (1988); Higgins *et al.* (1989) *CABIOS 5*:151-153; Corpet *et al.* (1988) *Nucleic Acids Res. 16*:10881-90; Huang *et al.* (1992) *Computer Applications in the Biosciences 8*:155-65, and Person *et al.* (1994) *Meth. Mol. Biol. 24*:307-331; preferred computer alignment methods also include the BLASTP, BLASTN, and BLASTX algorithms (see Altschul *et al.* (1990) *J. Mol. Biol. 215*:403-410). Alignment is also often performed by inspection and manual alignment. Sequence alignments are performed using the default parameters of the alignment programs.

(c) As used herein, "sequence identity" or "identity" in the context of two nucleic acid or polypeptide sequences makes reference to the residues in the two sequences that are the same when aligned for maximum correspondence over a specified comparison window. When percentage of sequence identity is used in reference to proteins it is recognized that residue positions which are not identical often differ by conservative amino acid substitutions, where amino acid residues are substituted for other amino acid residues with similar chemical properties (e.g., charge or hydrophobicity) and therefore do not change the functional properties of the molecule. When sequences differ in conservative substitutions, the percent sequence identity may be adjusted upwards to correct for the conservative nature of the substitution. Sequences that differ by such conservative substitutions are said to have "sequence similarity" or "similarity". Means for making this adjustment are well known to those of skill in the art. Typically this involves scoring a conservative substitution as a partial rather than a full mismatch, thereby increasing the percentage sequence identity. Thus, for example, where an identical amino acid is given a score of 1 and a non-conservative substitution is given a score of zero, a conservative substitution is given a score between zero and 1. The scoring of conservative substitutions is calculated, e.g., as implemented in the program PC/GENE (Intelligenetics, Mountain View, California).

(d) As used herein, "percentage of sequence identity" means the value determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the polynucleotide sequence in the comparison window may comprise additions or deletions (i.e., gaps) as compared to the

reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid base or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison, and multiplying the result by 100 to yield the percentage of sequence identity.

(e)(i) The term "substantial identity" of polynucleotide sequences means that a polynucleotide comprises a sequence that has at least 70% sequence identity, preferably at least 80%, more preferably at least 90%, and most preferably at least 95%, compared to a reference sequence using one of the alignment programs described using standard parameters. One of skill in the art will recognize that these values can be appropriately adjusted to determine corresponding identity of proteins encoded by two nucleotide sequences by taking into account codon degeneracy, amino acid similarity, reading frame positioning, and the like. Substantial identity of amino acid sequences for these purposes normally means sequence identity of at least 60%, more preferably at least 70%, 80%, 90%, and most preferably at least 95%.

Another indication that nucleotide sequences are substantially identical is if two molecules hybridize to each other under stringent conditions. Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. However, stringent conditions encompass temperatures in the range of about 1°C to about 20°C, depending upon the desired degree of stringency as otherwise qualified herein. Nucleic acids that do not hybridize to each other under stringent conditions are still substantially identical if the polypeptides they encode are substantially identical. This may occur, e.g., when a copy of a nucleic acid is created using the maximum codon degeneracy permitted by the genetic code. One indication that two nucleic acid sequences are substantially identical is when the polypeptide encoded by the first nucleic acid is immunologically cross reactive with the polypeptide encoded by the second nucleic acid.

(e)(ii) The term "substantial identity" in the context of a peptide indicates that a peptide comprises a sequence with at least 70% sequence identity to a

reference sequence, preferably 80%, more preferably 85%, most preferably at least 90% or 95% sequence identity to the reference sequence over a specified comparison window. Preferably, optimal alignment is conducted using the homology alignment algorithm of Needleman *et al.* (1970) *J. Mol. Biol. 48*:443. An indication that two peptide sequences are substantially identical is that one peptide is immunologically reactive with antibodies raised against the second peptide. Thus, a peptide is substantially identical to a second peptide, for example, where the two peptides differ only by a conservative substitution. Peptides that are "substantially similar" share sequences as noted above except that residue positions that are not identical may differ by conservative amino acid changes.

As described in detail below, the nucleotide sequences of the present invention may be operably linked to the nucleotide sequences encoding a protein of interest and thereby modulate the cellular localization of the protein. By "modulate" is intended any increase in the concentration of said protein in a plastid or compartment thereof beyond that which occurs in the absence of the transit peptide. Additionally, "modulate" refers to an increased rate of importation of said protein into the plastid as compared to the rate of importation in the absence of the transit peptide sequence.

The nucleotide sequences of the invention are provided in expression cassettes for expression in the plant of interest. The cassette will comprise a trancriptional initiation and translational termination sequence functional in plants operably linked to a nucleic acid sequence encoding a transit peptide of the invention, operably linked to a nucleotide encoding a protein of interest. The cassette may contain at least one additional sequence to be cotransformed into the organism. Alternatively, the additional sequences can be provided on another expression cassette.

"Operably linked" refers to a functional linkage between a promoter and a second sequence, wherein the promoter inititates and mediates transcription of DNA sequences corresponding to the second sequence. "Operably linked" also refers to a functional linkage between 2 or more distinct nucleotide sequences such that the nucleic acid sequences being linked are contiguous and, where necessary to join two protein coding regions, contiguous and in the same reading frame. Operably linking the transit peptide-coding sequences with the nucleotide

sequences encoding a protein of interest may require the manipulation of one or more of the DNA sequences. For example, a convenient restriction site or a linker sequences that acts as a non-specific spacer that may permit better recognition of the amino-terminal transit sequence may be introduced.

Expression of the coding sequences of a protein of interest operably linked to sequences of the transit peptide produces a hybrid polypeptide, or so-called fusion protein. By "hybrid" polypeptide is intended the coding sequences for the transit peptide is foreign to the coding sequences for the protein of interest, and hence, the two coding sequences are not natively expressed as a polypeptide in the plant cell.

It is recognized that in addition to the transit peptide of the present invention, additional amino acids may be fused to the protein of interest to further influence the fate the protein. Techniques for making fusion proteins recombinantly are well known in the art.

The transcriptional initiation region, the promoter, may be native or analogous or foreign or heterologous to the plant host. Additionally, the promoter may be a natural sequence or alternatively a synthetic sequence. By foreign is intended that the transcriptional initiation region is not found in the native plant into which the transcriptional initiation region is introduced.

While it may be preferable to express the sequences using heterologous promoters, the native promoter sequences of either the gene of interest or the transit peptide sequence may be used. Such constructs would change expression levels of the gene of interest in the plant or plant cell. Thus, the phenotype of the plant or plant cell is altered.

It is recognized that a variety of promoters will be useful in the invention, the choice of which will depend in part upon the desired level of expression of the protein of interest. It is recognized that the levels of expression can be controlled to modulate the levels of expression in the plant cell. Constitutive and tissue specific promoters are of particular interest. Such constitutive promoters include, for example, the core promoter of the Rsyn7 (copending U.S. Application Serial No. 08/661,601); the core CaMV 35S promoter (Odell *et al.* (1985) *Nature 313*:810-812); rice actin (McElroy *et al.* (1990) *Plant Cell 2*:163-171); ubiquitin (Christensen *et al.* (1989) *Plant Mol. Biol. 12*:619-632 and Christensen *et al.*

(1992) *Plant Mol. Biol. 18*:675-689); pEMU (Last *et al.* (1991) *Theor. Appl. Genet. 81*:581-588); MAS (Velten *et al.* (1984) *EMBO J. 3*:2723-2730); ALS promoter (U.S. Application Serial No. 08/409,297), and the like. Other constitutive promoters include, for example, U.S. Patent Nos. 5,608,149; 5,608,144; 5,604,121; 5,569,597; 5,466,785; 5,399,680; 5,268,463; and 5,608,142.

Tissue-specific promoters can be utilized to target enhanced expression within a particular plant tissue. Tissue-specific promoters include Yamamoto *et al.* (1997) *Plant J. 12(2)*255-265; Kawamata *et al.* (1997) *Plant Cell Physiol. 38(7)*:792-803; Hansen *et al.* (1997) *Mol. Gen Genet. 254(3)*:337-343; Russell *et al.* (1997) *Transgenic Res. 6(2)*:157-168; Rinehart *et al.* (1996) *Plant Physiol. 112(3)*:1331-1341; Van Camp *et al.* (1996) *Plant Physiol. 112(2)*:525-535; Canevascini *et al.* (1996) *Plant Physiol. 112(2)*:513-524; Yamamoto *et al.* (1994) *Plant Cell Physiol. 35(5)*:773-778; Lam (1994) *Results Probl. Cell Differ. 20*:181-196; Orozco *et al.* (1993) *Plant Mol Biol. 23(6)*:1129-1138; Matsuoka *et al.* (1993) *Proc Natl. Acad. Sci. USA 90(20)*:9586-9590; and Guevara-Garcia *et al.* (1993) *Plant J. 4(3)*:495-505. Such promoters can be modified, if necessary, for weak expression.

Leaf-specific promoters are known in the art. See, for example, Yamamoto *et al.* (1997) *Plant J. 12(2)*:255-265; Kwon *et al.* (1994) *Plant Physiol. 105*:357-67; Yamamoto *et al.* (1994) *Plant Cell Physiol. 35(5)*:773-778; Gotor *et al.* (1993) *Plant J. 3*:509-18; Orozco *et al.* (1993) *Plant Mol. Biol. 23(6)*:1129-1138; and Matsuoka *et al.* (1993) *Proc. Natl. Acad. Sci. USA 90(20)*:9586-9590.

The termination region may be native with the transcriptional initiation region, may be native with the DNA sequence of interest, or may be derived from another source. Convenient termination regions are available from the Ti-plasmid of *A. tumefaciens*, such as the octopine synthase and nopaline synthase termination regions. See also Guerineau *et al.* (1991) *Mol. Gen. Genet. 262*:141-144; Proudfoot (1991) *Cell 64*:671-674; Sanfacon *et al.* (1991) *Genes Dev. 5*:141-149; Mogen *et al.* (1990) *Plant Cell. 2*:1261-1272; Munroe *et al.* (1990) *Gene 91*:151-158; Ballas *et al.* (1989) *Nucleic Acids Res. 17*:7891-7903; Joshi *et al.* (1987) *Nucleic Acids Res. 15*:9627-9639.

The expression cassette contains a plurality of restriction sites to insert both the gene of interest and the transit peptide sequence 3' of the designated promoter.

The transit peptide sequences of the invention may be operably linked to the gene of interest at the 3' terminus, 5' terminus or internally. Preferably, the sequences of the invention will be placed at the 5' end. The nucleic acids included in the expression cassette may be optimized for expression in a plastid or compartment thereof to account for differences in codon usage between the plant nucleus and this organelle. In this manner, the nucleic acid sequences may be synthesized using chloroplast-preferred codons. See, for example, U.S. Patent No. 5,380,831, herein incorporated by reference.

Preferably the gene of interest encoding a protein to be localized the a plastid or compartment thereof is linked to the transit peptide nucleic acid sequence in such a way that upon translation and import into a plastid or compartment thereof the transit peptide is cleaved from the protein of interest. Methods for preparing transit peptide chimeras are known in the art and are described in the following publications and issued patents. See de Castro Silva Filho *et al.* (1996) *Plant Mol. Biol. 30*:769-780; Pilon *et al.* (1995) *J. Biol. Chem. 270(8)*:3882-3893; U.S. Patent No. 5,633,444; U.S. Patent No. 5,498,544.

The gene of interest may be native or analogous or foreign or heterologous to the plant host. By foreign is intended that the gene of interest is not found in the native plant into which it is introduced. The gene of interest may also be nuclear encoded or plastid encoded. Generally, the proteins selected for targeting to the plastids are heterologous to the transformed cell and nuclear encoded.

Genes of interest include, for example, any protein whose localization in the plastid will modify agronomically important traits such as oil, starch, and protein content. Other modified traits include herbicide, disease, and insect resistance.

Specific genes of interest may include, but are not limited to, the small subunit of ribulose bisphosphate carboxylase (Rubisco), Schnell *et al.* (1991) *J. Biol. Chem. 266(5)*:3335-3342; ferrodoxin, Pilon *et al.* (1995) *J. Biol. Chem. 270(8)*:3882-3893; light harvesting chlorophyll a/b binding protein, Lamppa *et al.* (1988) *J. Biol. Chem. 263*:14996-14999, Reiski FeS protein, Madueno *et al.* (1994) *J. Biol. Chem. 269(26)*:17458-17463; plastocyanin , Lawrence *et al.* (1997) *J. Biol. Chem. 272(33)*: 20357-20363; Bt1 protein, Li *et al.* (1992) *J. Biol. Chem. 267(26)*:18999-19004; forms of dihydropteroate synthase (DHPS), U.S. Patent No.

5,633,444; acetyl CoA carboxylase, U.S. Patent No. 5,498,544; superoxide dismutase, U.S. Patent No. 5,538,878; and 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS), Archer *et al.* (1990) *J. Bioenerg. Biomem. 22(6)*:789-810; U.S. Patent No. 5,188,642.

Of particular interest are genes encoding proteins involved in herbicide resistance. In a preferred embodiment, the herbicide resistance is imparted by 5-enolpyruvylshikimate-3-phosphate synthase. It is recognized that fragments and variants of the various proteins of interest may also be used with the transit peptide sequences of the invention. For example, the 5-enolpyruvylshikimate-3-phosphate synthase may be altered such that the protein product is less sensitive to herbicide inhibition. See for example, U.S. Patent No. 5,188,642.

Alternatively, the gene of interest may be a reporter gene. Reporter genes are generally known in the art. The reporter gene used should not be expressed endogenously. Ideally the reporter gene will exhibit low background activity and should not interfere with plant biochemical and physiological activities. The products expressed by the reporter gene should be stable and readily detectable. It is important that the reporter gene expression should be able to be assayed by a non-destructive, quantitative, sensitive, easy to perform and inexpensive method. Examples of suitable reporter genes known in the art can be found in, for example, Jefferson *et al.* (1991) in *Plant Molecular Biology Manual* (Gelvin *et al.* eds.) pp. 1-33, Kluwer Academic Publishers; DeWet *et al.* (1987) *Mol. Cell. Biol. 7*:725-737; Goff *et al.* (1990) *EMBO J. 9*:2517-2522; Kain *et al.* (1995) *BioTechniques 19*:650-655; Chiu *et al.* (1996) *Current Biology 6*:325-330.

The transit peptide sequences of the invention may be native or analogous or foreign or heterologous to either the host plant or to the gene of interest. By foreign is intended that the transit peptide sequence is not found in the native host plant or is not naturally encoded by the gene of interest. Furthermore, the DNA sequence encoding the transit peptide may be chemically synthesized either wholly or in part from the known sequence of the transit peptide.

Where appropriate, the gene(s) of interest and the transit peptide sequences of the invention may be optimized for increased expression in the transformed plant. That is, the genes can be synthesized using plant-preferred codons for improved expression. See, for example, Campbell and Gowri (1990) *Plant Physiol.*

92:1-11 for a discussion of host-preferred codon usage. Methods are available in the art for synthesizing plant-preferred genes. See, for example, U.S. Patent Nos. 5,380,831, 5,436,391, and Murray *et al.* (1989) *Nucleic Acids Res. 17*:477-498, herein incorporated by reference.

Additional sequence modifications are known to enhance gene expression in a cellular host. These include elimination of sequences encoding spurious polyadenylation signals, exon-intron splice site signals, transposon-like repeats, and other such well-characterized sequences that may be deleterious to gene expression. The G-C content of the sequence may be adjusted to levels average for a given cellular host, as calculated by reference to known genes expressed in the host cell. When possible, the sequence is modified to avoid predicted hairpin secondary mRNA structures.

The expression cassettes may additionally contain 5' leader sequences in the expression cassette construct. Such leader sequences can act to enhance translation. Translation leaders are known in the art and include: picornavirus leaders, for example, EMCV leader (Encephalomyocarditis 5' noncoding region) (Elroy-Stein *et al.* (1989) *PNAS USA 86*:6126-6130); potyvirus leaders, for example, TEV leader (Tobacco Etch Virus) (Allison *et al.* (1986); MDMV leader (Maize Dwarf Mosaic Virus); *Virology 154*:9-20), and human immunoglobulin heavy-chain binding protein (BiP), (Macejak *et al.* (1991) *Nature 353*:90-94); untranslated leader from the coat protein mRNA of alfalfa mosaic virus (AMV RNA 4) (Jobling *et al.* (1987) *Nature 325*:622-625); tobacco mosaic virus leader (TMV) (Gallie *et al.* (1989) in *Molecular Biology of RNA*, ed. Cech (Liss, New York), pp. 237-256); and maize chlorotic mottle virus leader (MCMV) (Lommel *et al.* (1991) *Virology 81*:382-385). See also, Della-Cioppa *et al.* (1987) *Plant Physiol. 84*:965-968. Other methods known to enhance translation can also be utilized, for example, introns, and the like.

In preparing the expression cassette, the various DNA fragments may be manipulated, so as to provide for the DNA sequences in the proper orientation and, as appropriate, in the proper reading frame. Toward this end, adapters or linkers may be employed to join the DNA fragments or other manipulations may be involved to provide for convenient restriction sites, removal of superfluous DNA, removal of restriction sites, or the like. For this purpose, *in vitro* mutagenesis,

primer repair, restriction, annealing, resubstitutions, *e.g.*, transitions and transversions, may be involved.

Generally, the expression cassette will comprise a selectable marker gene for the selection of transformed cells. Selectable marker genes are utilized for the selection of transformed cells or tissues. Marker genes include genes encoding antibiotic resistance, such as those encoding neomycin phosphotransferase II (NEO) and hygromycin phosphotransferase (HPT), as well as genes conferring resistance to herbicidal compounds, such as glufosinate ammonium, bromoxynil, imidazolinones, and 2,4-dichlorophenoxyacetate (2,4-D). See generally, Yarranton (1992) *Curr. Opin. Biotech. 3*:506-511; Christopherson *et al.* (1992) *Proc. Natl. Acad. Sci. USA 89*:6314-6318; Yao *et al.* (1992) *Cell 71*:63-72; Reznikoff (1992) *Mol. Microbiol. 6*:2419-2422; Barkley *et al.* (1980) in *The Operon*, pp. 177-220; Hu *et al.* (1987) *Cell 48*:555-566; Brown *et al.* (1987) *Cell 49*:603-612; Figge *et al.* (1988) *Cell 52*:713-722; Deuschle *et al.* (1989) *Proc. Natl. Acad. Aci. USA 86*:5400-5404; Fuerst *et al.* (1989) *Proc. Natl. Acad. Sci. USA 86*:2549-2553; Deuschle *et al.* (1990) *Science 248*:480-483; Gossen (1993) Ph.D. Thesis, University of Heidelberg; Reines *et al.* (1993) *Proc. Natl. Acad. Sci. USA 90*:1917-1921; Labow *et al.* (1990) *Mol. Cell. Biol. 10*:3343-3356; Zambretti *et al.* (1992) *Proc. Natl. Acad. Sci. USA 89*:3952-3956; Baim *et al.* (1991) *Proc. Natl. Acad. Sci. USA 88*:5072-5076; Wyborski *et al.* (1991) *Nucleic Acids Res. 19*:4647-4653; Hillenand-Wissman (1989) *Topics Mol. Struc. Biol. 10*:143-162; Degenkolb *et al.* (1991) *Antimicrob. Agents Chemother. 35*:1591-1595; Kleinschnidt *et al.* (1988) *Biochemistry 27*:1094-1104; Bonin (1993) Ph.D. Thesis, University of Heidelberg; Gossen *et al.* (1992) *Proc. Natl. Acad. Sci. USA 89*:5547-5551; Oliva *et al.* (1992) *Antimicrob. Agents Chemother. 36*:913-919; Hlavka *et al.* (1985) *Handbook of Experimental Pharmacology*, Vol. 78 (Springer-Verlag, Berlin); Gill *et al.* (1988) *Nature 334*:721-724. Such disclosures are herein incorporated by reference.

The above list of selectable marker genes is not meant to be limiting. Any selectable marker gene can be used in the present invention.

The present invention also relates to the introduction of the transformation constructs into plant protoplasts, calli, tissues, or organ explants and the regeneration of transformed plants expressing the recombinant constructs of the invention.

The expression cassette sequences of the present invention may be used for transformation of any plant species, including, but not limited to, corn (*Zea mays*), canola (*Brassica napus, Brassica rapa* ssp.), alfalfa (*Medicago sativa*), rice (*Oryza sativa*), rye (*Secale cereale*), sorghum (*Sorghum bicolor, Sorghum vulgare*), sunflower (*Helianthus annuus*), wheat (*Triticum aestivum*), soybean (*Glycine max*), tobacco (*Nicotiana tabacum*), potato (*Solanum tuberosum*), peanuts (*Arachis hypogaea*), cotton (*Gossypium barbadense, Gossypium hirsutum*), sweet potato (*Ipomoea batatus*), cassava (*Manihot esculenta*), coffee (*Cofea* spp.), coconut (*Cocos nucifera*), pineapple (*Ananas comosus*), citrus trees (*Citrus* spp.), cocoa (*Theobroma cacao*), tea (*Camellia sinensis*), banana (*Musa* spp.), avocado (*Persea americana*), fig (*Ficus casica*), guava (*Psidium guajava*), mango (*Mangifera indica*), olive (*Olea europaea*), papaya (*Carica papaya*), cashew (*Anacardium occidentale*), macadamia (*Macadamia integrifolia*), almond (*Prunus amygdalus*), sugar beets (*Beta vulgaris*), sugarcane (*Saccharum* spp.), oats, barley, vegetables, ornamentals, and conifers.

Vegetables include tomatoes (*Lycopersicon esculentum*), lettuce (e.g., *Lactuca sativa*), green beans (*Phaseolus vulgaris*), lima beans (*Phaseolus limensis*), peas (*Lathyrus* spp.), and members of the genus *Cucumis* such as cucumber (*C. sativus*), cantaloupe (*C. cantalupensis*), and musk melon (*C. melo*). Ornamentals include azalea (*Rhododendron* spp.), hydrangea (*Macrophylla hydrangea*), hibiscus (*Hibiscus rosasanensis*), roses (*Rosa* spp.), tulips (*Tulipa* spp.), daffodils (*Narcissus* spp.), petunias (*Petunia hybrida*), carnation (*Dianthus caryophyllus*), poinsettia (*Euphorbia pulcherrima*), and chrysanthemum. Conifers that may be employed in practicing the present invention include, for example, pines such as loblolly pine (*Pinus taeda*), slash pine (*Pinus elliotii*), ponderosa pine (*Pinus ponderosa*), lodgepole pine (*Pinus contorta*), and Monterey pine (*Pinus radiata*); Douglas-fir (*Pseudotsuga menziesii*); Western hemlock (*Tsuga canadensis*); Sitka spruce (*Picea glauca*); redwood (*Sequoia sempervirens*); true firs such as silver fir (*Abies amabilis*) and balsam fir (*Abies balsamea*); and cedars such as Western red cedar (*Thuja plicata*) and Alaska yellow-cedar (*Chamaecyparis nootkatensis*). Preferably, plants of the present invention are crop plants (for example, corn, alfalfa, sunflower, canola, soybean, cotton, peanut, sorghum, wheat, tobacco, etc.), more preferably corn and soybean plants, yet more preferably corn plants.

Transformation protocols as well as protocols for introducing nucleotide sequences into plants may vary depending on the type of plant or plant cell, i.e., monocot or dicot, targeted for transformation. Suitable methods of introducing nucleotide sequences into plant cells and subsequent insertion into the plant genome include microinjection (Crossway *et al.* (1986) *Biotechniques 4*:320-334), electroporation (Riggs *et al.* (1986) *Proc. Natl. Acad. Sci. USA 83*:5602-5606, *Agrobacterium*-mediated transformation (Townsend *et al.*, U.S. Pat No. 5,563,055), direct gene transfer (Paszkowski *et al.* (1984) *EMBO J. 3*:2717-2722), and ballistic particle acceleration (see, for example, Sanford *et al.*, U.S. Patent No. 4,945,050; Tomes *et al.* (1995) "Direct DNA Transfer into Intact Plant Cells via Microprojectile Bombardment," in *Plant Cell, Tissue, and Organ Culture: Fundamental Methods*, ed. Gamborg and Phillips (Springer-Verlag, Berlin); and McCabe *et al.* (1988) *Biotechnology 6*:923-926). Also see Weissinger *et al.* (1988) *Ann. Rev. Genet. 22*:421-477; Sanford *et al.* (1987) *Particulate Science and Technology 5*:27-37 (onion); Christou *et al.* (1988) *Plant Physiol. 87*:671-674 (soybean); McCabe *et al.* (1988) *Bio/Technology 6*:923-926 (soybean); Finer and McMullen (1991) *In Vitro Cell Dev. Biol. 27P*:175-182 (soybean); Singh *et al.* (1998) *Theor. Appl. Genet. 96*:319-324 (soybean); Datta *et al.* (1990) *Biotechnology 8*:736-740 (rice); Klein *et al.* (1988) *Proc. Natl. Acad. Sci. USA 85*:4305-4309 (maize); Klein *et al.* (1988) *Biotechnology 6*:559-563 (maize); Tomes, U.S. Patent No. 5,240,855; Buising *et al.*, U.S. Patent Nos. 5,322,783 and 5,324,646; Tomes *et al. (1995)* "Direct DNA Transfer into Intact Plant Cells via Microprojectile Bombardment," in *Plant Cell, Tissue, and Organ Culture: Fundamental Methods*, ed. Gamborg (Springer-Verlag, Berlin) (maize); Klein *et al.* (1988) *Plant Physiol. 91*:440-444 (maize); Fromm *et al.* (1990) *Biotechnology 8*:833-839 (maize); Hooykaas-Van Slogteren *et al.* (1984) *Nature (London) 311*:763-764; Bowen *et al.*, U.S. Patent No. 5,736,369 (cereals); Bytebier *et al.* (1987) *Proc. Natl. Acad. Sci. USA 84*:5345-5349 (Liliaceae); De Wet *et al.* (1985) in *The Experimental Manipulation of Ovule Tissues*, ed. Chapman *et al.* (Longman, New York), pp. 197-209 (pollen); Kaeppler *et al.* (1990) *Plant Cell Reports 9*:415-418 and Kaeppler *et al.* (1992) *Theor. Appl. Genet. 84*:560-566 (whisker-mediated transformation); D'Halluin *et al.* (1992) *Plant Cell 4*:1495-1505 (electroporation); Li *et al.* (1993) *Plant Cell Reports 12*:250-255 and Christou and Ford (1995)

Annals of Botany 75:407-413 (rice); Osjoda *et al.* (1996) *Nature Biotechnology 14*:745-750 (maize via *Agrobacterium tumefaciens*); all of which are herein incorporated by reference.

The modified plant may be grown into plants in accordance with conventional ways. See, for example, McCormick *et al.* (1986) *Plant Cell. Reports 5*:81-84. These plants may then be grown, and either pollinated with the same transformed strain or different strains, and the resulting hybrid having the desired phenotypic characteristic identified. Two or more generations may be grown to ensure that the subject phenotypic characteristic is stably maintained and inherited and then seeds harvested to ensure the desired phenotype or other property has been achieved.

Assays to determine the efficiency by which the isolated transit peptide sequences of the invention target a protein of interest to a plastid are known. A reporter gene such as β-glucuronidase (GUS), chloramphenicol acetyl transferase (CAT), or green fluorescent protein (GFP) is operably linked to the transit peptide sequence. This fusion is placed behind the control of a suitable promoter, ligated into a transformation vector, and transformed into a plant or plant cell. Following an adequate period of time for expression and localization into the plastid, the plastid fraction is extracted and reporter activity assayed. The ability of the isolated sequences to target and deliver the reporter protein to the plastid will be compared to other known transit peptide sequences. See de Castro Silva Filho *et al.* (1996) *Plant Mol. Biol. 30*:769-780. Protein import can also be verified *in vitro* through the addition of proteases to the isolated plastid fraction. Proteins which were successfully imported into the plastid are resistant to the externally added proteases whereas proteins that remain in the cytosol are susceptible to digestion.

The following examples are offered by way of illustration and not by way of limitation.

EXPERIMENTAL

Example 1

Generating an expression cassette for the expression of GFP-transit peptide chimeric proteins

The Green Fluorescent Protein (GFP) gene described by Prasher *et al.* (1992) *Gene 111*:229-233, is modified for expression in maize. The modified sequence, GFPm, is derived from a back translation of the GFP protein sequence using maize preferred codons and is shown in SEQ ID NO:23. Sequence analysis is performed using the Wisconsin Sequence Analysis Package from Genetics Computer Group, Madison, WI. The nucleotide sequence is assembled from a series of synthetic oligonucleotides. Cloning sites within the GFPm include a 5' flanking *Bam*HI restriction site, an *Afl*III site at the start codon, a 3' flanking *Hpa*I site or a *Bgl*II site converting the stop codon to an isoleucine.

Amino terminal and carboxy terminal fusions of transit peptide sequences to the modified green fluorescent protein (GFPm) are created using synthetic oligonucleotides encoding the transit peptide sequence flanked by appropriate restriction sites that allow in-frame fusions with GFPm.

The ubiquitin promoter is inserted upstream of the GFPm fusion protein. Also engineered into the expression cassette is an intron and a PinII termination sequence.

Example 2

Transformation and Regeneration of Transgenic Plants with GFP Screening

Immature maize embryos from greenhouse donor plants are bombarded with a plasmid containing a transit peptide sequence cloned into a flanking restriction sites of GFPm to create an inframe fusion. These sequences are operably linked to a ubiquitin promoter (Figure 1). Also contained on this plasmid is the selectable marker gene, PAT, (Wohlleben *et al.* (1988) *Gene 70*:25-37) that confers resistance to the herbicide Bialaphos. Transformation is performed as follows. All media recipes are in the Appendix.

Preparation of Target Tissue

The ears are surface sterilized in 30% Chlorox bleach plus 0.5% Micro detergent for 20 minutes, and rinsed two times with sterile water. The immature embryos are excised and placed embryo axis side down (scutellum side up), 25 embryos per plate, on 560Y medium for 4 hours and then aligned within the 2.5-cm target zone in preparation for bombardment.

Preparation of DNA

A plasmid vector comprising a transit peptide sequence sequences cloned into restriction sites resulting in an inframe fusion with GFPm, and operably linked to a ubiquitin promoter, and containing a PAT selectable marker is precipitated onto 1.1 μm (average diameter) tungsten pellets using a $CaCl_2$ precipitation procedure as follows:

100 μl prepared tungsten particles in water
10 μl (1 μg) DNA in TrisEDTA buffer (1 μg total)
100 μl 2.5 M $CaCl_2$
10 μl 0.1 M spermidine

Each reagent is added sequentially to the tungsten particle suspension, while maintained on the multitube vortexer. The final mixture is sonicated briefly and allowed to incubate under constant vortexing for 10 minutes. After the precipitation period, the tubes are centrifuged briefly, liquid removed, washed with 500 ml 100% ethanol, and centrifuged for 30 seconds. Again the liquid is removed, and 105 μl 100% ethanol is added to the final tungsten particle pellet. For particle gun bombardment, the tungsten/DNA particles are briefly sonicated and 10 μl spotted onto the center of each macrocarrier and allowed to dry about 2 minutes before bombardment.

Particle Gun Treatment

The sample plates are bombarded at level #4 in particle gun #HE34-1 or #HE34-2. All samples receive a single shot at 650 PSI, with a total of ten aliquots taken from each tube of prepared particles/DNA.

Subsequent Treatment

Following bombardment, the embryos are kept on 560Y medium for 2 days, then transferred to 560R selection medium containing 3 mg/liter Bialaphos, and subcultured every 2 weeks. After approximately 10 weeks of selection, selection-resistant callus clones are transferred to 288J medium to initiate plant regeneration. Following somatic embryo maturation (2-4 weeks), well-developed somatic embryos are transferred to medium for germination and transferred to the lighted culture room. Approximately 7-10 days later, developing plantlets are transferred to 272V hormone-free medium in tubes for 7-10 days until plantlets are well established. Plants are then transferred to inserts in flats (equivalent to 2.5" pot) containing potting soil and grown for 1 week in a growth chamber, subsequently grown an additional 1-2 weeks in the greenhouse, then transferred to classic 600 pots (1.6 gallon) and grown to maturity. Screening for GFP expression is carried out at each transfer using a Xenon and/or Mercury light source with the appropriate filters for GFP visualization.

Once GFP expressing colonies are identified they are monitored regularly for new growth and expression using the Xenon light source. Plant cells containing GFP are regenerated by transferring the callus to 288 medium containing MS salts, 1 mg/L IAA, 0.5 mg/L zeatin and 4% sucrose. The callus is placed in the light. As plantlets develop they are transferred to tubes containing 272K, hormone-free MS medium and 3% sucrose. The percentage of green fluorescent colonies that regenerated into whole plants can be determined.

The ability of the tansit peptide to target GFP to the plastid is determined in stable transgenic maize cells using epifluorescent microscopy and image enhancement software. Samples of calli from the transformed maize plants are fixed in FAA and are examined with UV filters to visualize GFP localization in the plastid.

APPENDIX

272 V

Ingredient	Amount	Unit
D-I H_2O	950.000	Ml
MS Salts (GIBCO 11117-074)	4.300	G
Myo-Inositol	0.100	G
MS Vitamins Stock Solution ##	5.000	Ml
Sucrose	40.000	G
Bacto-Agar @	6.000	G

Directions:

@ = Add after bringing up to volume

Dissolve ingredients in polished D-I H_2O in sequence

Adjust to pH 5.6

Bring up to volume with polished D-I H_2O after adjusting pH

Sterilize and cool to 60°C.

= Dissolve 0.100 g of Nicotinic Acid; 0.020 g of Thiamine.HCL; 0.100 g of Pyridoxine.HCL; and 0.400 g of Glycine in 875.00 ml of polished D-I H_2O in sequence. Bring up to volume with polished D-I H_2O. Make in 400 ml portions. Thiamine.HCL & Pyridoxine.HCL are in Dark Desiccator. Store for one month, unless contamination or precipitation occurs, then make fresh stock.

Total Volume (L) = 1.00

288 J

Ingredient	Amount	Unit
D-I H_2O	950.000	Ml
MS Salts	4.300	g
Myo-Inositol	0.100	g
MS Vitamins Stock Solution ##	5.000	ml
Zeatin .5mg/ml	1.000	ml
Sucrose	60.000	g
Gelrite @	3.000	g
Indoleacetic Acid 0.5 mg/ml #	2.000	ml
0.1mM Abscisic Acid	1.000	ml
Bialaphos 1mg/ml #	3.000	ml

Directions:

@ = Add after bringing up to volume

Dissolve ingredients in polished D-I H_2O in sequence

Adjust to pH 5.6

Bring up to volume with polished D-I H_2O after adjusting pH

Sterilize and cool to 60°C.

Add 3.5g/L of Gelrite for cell biology.

= Dissolve 0.100 g of Nicotinic Acid; 0.020 g of Thiamine.HCL; 0.100 g of Pyridoxine.HCL; and 0.400 g of Glycine in 875.00 ml of polished D-I H_2O in sequence. Bring up to volume with polished D-I H_2O. Make in 400 ml portions. Thiamine.HCL & Pyridoxine.HCL are in Dark Desiccator. Store for one month, unless contamination or precipitation occurs, then make fresh stock.

Total Volume (L) = 1.00

560 R

Ingredient	Amount	Unit
D-I Water, Filtered	950.000	ml
CHU (N6) Basal Salts (SIGMA C-1416)	4.000	g
Eriksson's Vitamin Mix (1000X SIGMA-1511)	1.000	ml
Thiamine.HCL 0.4mg/ml	1.250	ml
Sucrose	30.000	g
2, 4-D 0.5mg/ml	4.000	ml
Gelrite @	3.000	g
Silver Nitrate 2mg/ml #	0.425	ml
Bialaphos 1mg/ml #	3.000	ml

Directions:

@ = Add after bringing up to volume

= Add after sterilizing and cooling to temp.

Dissolve ingredients in D-I H_2O in sequence

Adjust to pH 5.8 with KOH

Bring up to volume with D-I H_2O

Sterilize and cool to room temp.

Total Volume (L) = 1.00

560 Y

Ingredient	Amount	Unit
D-I Water, Filtered	950.000	ml
CHU (N6) Basal Salts (SIGMA C-1416)	4.000	g
Eriksson's Vitamin Mix (1000X SIGMA-1511)	1.000	ml
Thiamine.HCL 0.4mg/ml	1.250	ml
Sucrose	120.000	g
2,4-D 0.5mg/ml	2.000	ml
L-Proline	2.880	g
Gelrite @	2.000	g
Silver Nitrate 2mg/ml #	4.250	ml

Directions:

@ = Add after bringing up to volume

= Add after sterilizing and cooling to temp.

Dissolve ingredients in D-I H_2O in sequence

Adjust to pH 5.8 with KOH

Bring up to volume with D-I H_2O

Sterilize and cool to room temp.

** Autoclave less time because of increased sucrose**

Total Volume (L) = 1.00

All publications and patent applications mentioned in the specification are indicative of the level of those skilled in the art to which this invention pertains. All publications and patent applications are herein incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be obvious that certain changes and modifications may be practiced within the scope of the appended claims.

THAT WHICH IS CLAIMED

1. A method of modulating the subcellular localization of a protein of interest in a plant or plant cell, said method comprising transforming said plant or plant cell with an expression cassette comprising a promoter operably linked to a nucleotide sequence encoding a transit peptide operably linked to a nucleotide sequence encoding a protein of interest, wherein said transit peptide directs the protein of interest to a plant plastid and said transit peptide is selected from the group consisting of:

a) a nucleic acid molecule encoding a polypeptide comprising an amino acid sequence set forth in one of the SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, or 22;

b) a nucleic acid molecule comprising a sequence set forth in one of the SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, or 21;

c) a nucleic acid molecule hybridizing under stringent conditions to the sequences of a) or b).

2. The method of claim 1, wherein said plant plastid is selected from the group comprising a chloroplast, amyloplast, chromoplast, and leucoplast.

3. The method of claim 1, wherein said promoter is a constitutive promoter.

4. The method of claim 1, wherein said promoter is a tissue-specific promoter.

5. The method of claim 1, wherein said protein of interest imparts herbicide resistance.

6. The method of claim 5, wherein said protein of interest is 5-enolpyruvylshikimate-3-phosphate synthase.

7. An isolated nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

a) a nucleotide sequence encoding a polypeptide comprising the amino acid sequence set forth in one of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, or 22;

b) a nucleotide sequence comprising the sequence set forth in SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, or 21;

c) a nucleotide sequence hybridizing under stringent conditions to a nucleotide sequence of a) or b).

8. An expression cassette comprising a promoter operably linked to a sequence encoding a transit peptide operably linked to a gene of interest, wherein said sequence encoding a transit peptide is selected from the group consisting of:

a) a nucleic acid molecule encoding a polypeptide comprising an amino acid sequence set forth in one of the SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, or 22;

b) a nucleic acid sequence comprising a sequence set forth in one of the SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, or 21;

c) a nucleic acid sequence hybridizing under stringent conditions to a sequence of a) or b).

9. A vector comprising the expression cassette of claim 8.

10. A transformed plant having stably incorporated in its genome an expression cassette comprising the following operably linked elements; a promoter, a coding sequence for a protein of interest, and a nucleotide sequence encoding a transit peptide, wherein said nucleotide sequence encoding a transit peptide is selected from the group consisting of:

a) a nucleic acid molecule encoding a polypeptide comprising an amino acid sequence set forth in one of the SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, or 22;

b) a nucleic acid molecule comprising a sequence set forth in one of the SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, or 21;

c) a nucleic acid sequence hybridizing under stringent conditions to the sequences of a) or b).

11. The plant of claim 10, wherein said plant is a dicot.

12. The plant of claim 10, wherein said plant is a monocot.

13. The plant of claim 12, wherein said monocot is maize.

14. Seed of the plant of claim 10.

15. A transformed plant cell having stably incorporated in its genome an expression cassette comprising the following operably linked elements; a promoter, a coding sequence for a protein of interest, a nucleotide sequence encoding a transit peptide, wherein said sequence encoding the transit peptide is selected from the group consisting of:

a) a nucleic acid molecule encoding a polypeptide comprising an amino acid sequence set forth in one of the SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, or 22;

b) a nucleic acid molecule comprising a sequence set forth in one of the SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, or 21;

c) a nucleic acid molecule hybridizing under stringent conditions to the sequences of a) or b).

16. An isolated polypeptide selected from the group consisting of:

a) a polypeptide comprising an amino acid sequence set forth in SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, or 22;

b) a polypeptide encoded by a nucleotide sequence comprising the sequence set forth in SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, or 21;

c) a polypeptide encoded by a nucleotide sequence that hybridizes under stringent conditions to a nucleotide sequence comprising the sequence set forth in SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, or 21.

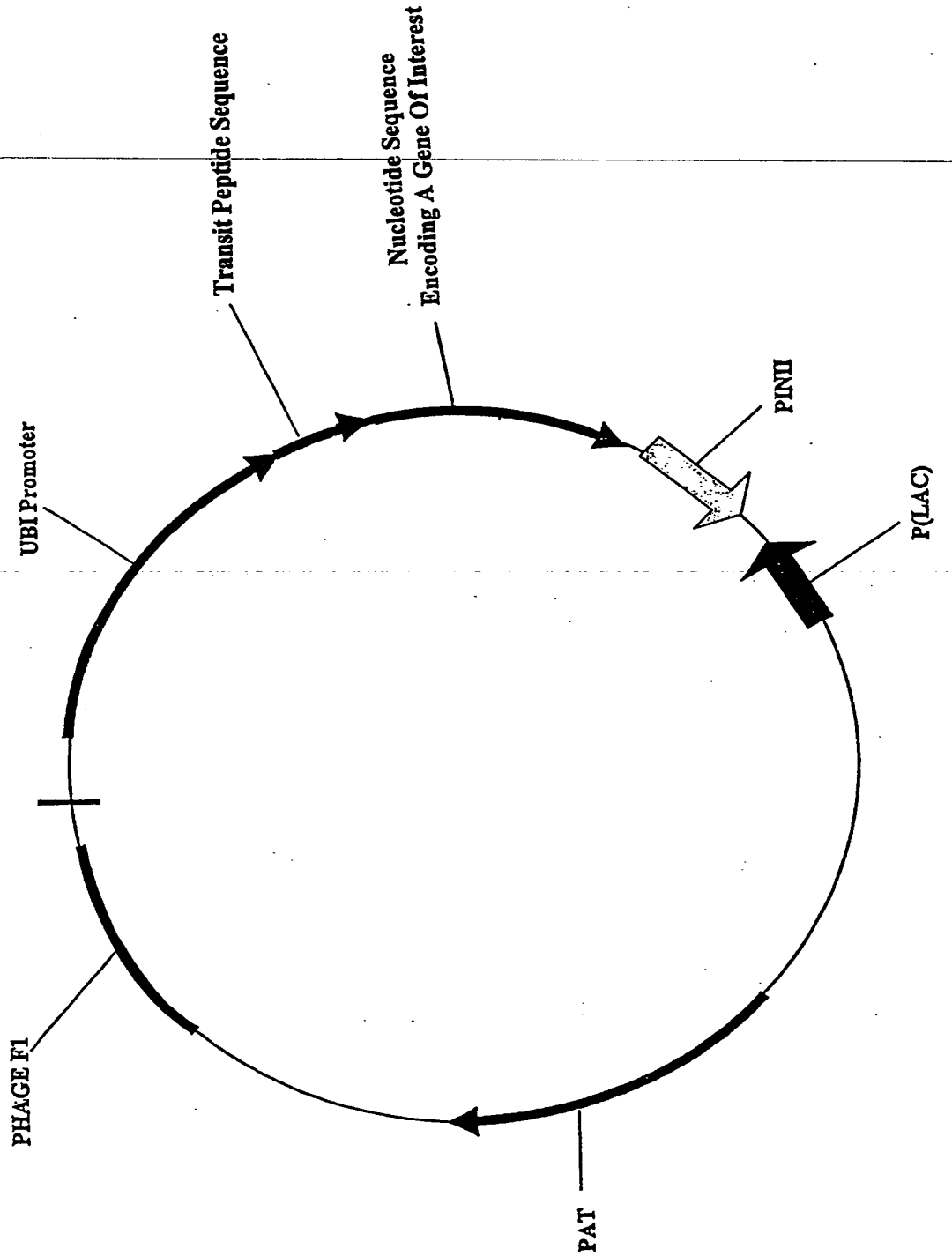

Transit Peptide Sequence
Nucleotide Sequence
Encoding A Gene Of Interest
PINII
P(LAC)
UBI Promoter
PHAGE F1
PAT

Figure 1

SEQUENCE LISTING

<110> Benson, Robert J.

<120> Organelle Targeting sequences

<130> 5718-30

<160> 24

<170> FastSEQ for Windows Version 3.0

<210> 1
<211> 407
<212> DNA
<213> Zea mays

<220>
<221> misc_feature
<222> (1)...(407)
<223> n = A,T,C or G

<400> 1
ccggtccaag gnntttcacg ctcagcttca gcaatggctg cctccaccat ggcgatctcc 60
tccacggcga tggccggcac ccccatcaag gtgggttcct tcggcgaggg ccgcatcacc 120
atgcgcaaga ccgtgggcaa gcccaaggtg gcggcgtccg gcagcccctg gtacggcccc 180
gaccgcgtca agtacctcgg ccccttctcc ggcgagcccc cgagctacct caccggcgag 240
ttccccggcg actacggctg ggacaccgcc gggctgtccg ccgaccccga gacattcgcc 300
aagaaccgcg agctggaggt gatccactcc cgctgggcca tgctcggcgc gctcggctgc 360
gtcttccccg agtgctctcc cgcaacggcg tcaagttcgg cgaggcg 407

<210> 2
<211> 135
<212> PRT
<213> Zea mays

<220>
<221> VARIANT
<222> (1)...(135)
<223> Xaa = Any Amino Acid

<400> 2
Pro Val Gln Gly Xaa Ser Arg Ser Ala Ser Ala Met Ala Ala Ser Thr
1 5 10 15
Met Ala Ile Ser Ser Thr Ala Met Ala Gly Thr Pro Ile Lys Val Gly
20 25 30
Ser Phe Gly Glu Gly Arg Ile Thr Met Arg Lys Thr Val Gly Lys Pro
35 40 45
Lys Val Ala Ala Ser Gly Ser Pro Trp Tyr Gly Pro Asp Arg Val Lys
50 55 60
Tyr Leu Gly Pro Phe Ser Gly Glu Pro Pro Ser Tyr Leu Thr Gly Glu
65 70 75 80
Phe Pro Gly Asp Tyr Gly Trp Asp Thr Ala Gly Leu Ser Ala Asp Pro
85 90 95
Glu Thr Phe Ala Lys Asn Arg Glu Leu Glu Val Ile His Ser Arg Trp

```
            100                 105                 110
Ala Met Leu Gly Ala Leu Gly Cys Val Phe Pro Glu Cys Ser Pro Ala
        115                 120                 125
Thr Ala Ser Ser Ser Ala Arg
    130                 135
```

<210> 3
<211> 426
<212> DNA
<213> Zea mays

<220>
<221> misc_feature
<222> (1)...(426)
<223> n = A,T,C or G

<400> 3

```
tttctnnncc cctcagcctc gccacttcag agagagagag agagagagaa gagagagaga      60
gagagagagc taacgatggc gtccagagac gcggccagca gccccaaatc agcgccgcgt     120
cgccatgggc cggcggcgcc caagaatggc gccgacctgc tcagcccacg gttccggtcg     180
gctgccgcgc tggccggctg ggacgaggag tccgtcctcc ttgccgcgat ggtcgtcgac     240
gacaccccag tccgcgactc ccgtcgcaag aggcgggcct ccaactcctc cgtcggcggc     300
agcgccggat tccagcacca ggaagaggag gccgcgcagg ctgtctccgg ccaagatccc     360
gccggtggtt ttcgcgctcg acgacgaaaa gtcggacact gcggcagata gcaaatcgga     420
attaag                                                               426
```

<210> 4
<211> 117
<212> PRT
<213> Zea mays

<400> 4

```
Arg Trp Arg Pro Glu Thr Arg Pro Ala Ala Pro Asn Gln Arg Arg Val
 1               5                  10                  15
Ala Met Gly Arg Arg Arg Pro Arg Met Ala Pro Thr Cys Ser Ala His
            20                  25                  30
Gly Ser Gly Arg Leu Pro Arg Trp Pro Ala Gly Thr Arg Ser Pro Ser
        35                  40                  45
Ser Leu Pro Arg Trp Ser Ser Thr Thr Pro Gln Ser Ala Thr Pro Val
    50                  55                  60
Ala Arg Gly Gly Pro Pro Thr Pro Pro Ser Ala Ala Ala Pro Asp Ser
65                  70                  75                  80
Ser Thr Arg Lys Arg Arg Pro Arg Arg Leu Ser Pro Ala Lys Ile Pro
                85                  90                  95
Pro Val Val Phe Ala Leu Asp Asp Glu Lys Ser Asp Thr Ala Ala Asp
            100                 105                 110
Ser Lys Ser Glu Leu
        115
```

<210> 5
<211> 455
<212> DNA
<213> Zea mays

<220>
<221> misc_feature
<222> (1)...(455)

<223> n = A,T,C or G

<400> 5
gcaactctaa tagatcatat aggaaacggt agcctgcagt accggtccgg aattcccggg 60
nnnnnnnnnn nnnnnggttg ataacctgtt gaacagtgag gatgtccact acatgctcgg 120
ggccttgagg actcttggtc tctctgtcga agcggacaaa gctgccaaaa gagctgtagt 180
tgttggctgt ggtggaaagt tcccagttga ggattctaaa gaggaagtgc agctcttctt 240
ggggaatgct ggaactgcaa tgcggccatt gacagcagct gttactgctg ctggtggaaa 300
tgcaacttac gtgctttatg gagtaccaag aatgagggag agacccattg gcgacttggt 360
tgtcggattn aaancacttg gtgcagatgt tgattgtttc cttggcactg aatgccccac 420
ctgttcgtgt caanggaatc gganggccna cctgg 455

<210> 6
<211> 144
<212> PRT
<213> Zea mays

<220>
<221> VARIANT
<222> (1)...(144)
<223> Xaa = Any Amino Acid

<400> 6
Glu Thr Val Ala Cys Ser Thr Gly Pro Glu Phe Pro Gly Xaa Xaa Xaa
1 5 10 15
Xaa Xaa Val Asp Asn Leu Leu Asn Ser Glu Asp Val His Tyr Met Leu
20 25 30
Gly Ala Leu Arg Thr Leu Gly Leu Ser Val Glu Ala Asp Lys Ala Ala
35 40 45
Lys Arg Ala Val Val Val Gly Cys Gly Gly Lys Phe Pro Val Glu Asp
50 55 60
Ser Lys Glu Glu Val Gln Leu Phe Leu Gly Asn Ala Gly Thr Ala Met
65 70 75 80
Arg Pro Leu Thr Ala Ala Val Thr Ala Ala Gly Gly Asn Ala Thr Tyr
85 90 95
Val Leu Tyr Gly Val Pro Arg Met Arg Glu Arg Pro Ile Gly Asp Leu
100 105 110
Val Val Gly Xaa Lys Xaa Leu Gly Ala Asp Val Asp Cys Phe Leu Gly
115 120 125
Thr Glu Cys Pro Thr Cys Ser Cys Gln Xaa Asn Arg Xaa Ala Xaa Leu
130 135 140

<210> 7
<211> 459
<212> DNA
<213> Zea mays

<220>
<221> misc_feature
<222> (1)...(459)
<223> n = A,T,C or G

<400> 7
cnnnnnnnnn nnngctgctc caaggaggcg gacagaggca ggcaggaagg cactcacttc 60
tccaggagcc cggagtcggc agcaagagag aatggccgcc gaggncgccg ccgccgccac 120
ccttcgtact gctctctccc aagcaacagg gcaagagcag agagcttcac tgctgtgcac 180
acctgcgggc gccagtgtga gcatcccgaa gcaatggtac aacctcatcg ccgacctgcc 240

ggtgaagcca ccgccaccgc tgcacccgca gacacaccag cctctgaatc ccagcgacct 300
ctcccctctg ttccccgacg agctgatcag gcaggaggtc accgacgagc ggttcgtcga 360
catacccgag gaggtcatcg acgtgtacaa gctctggcgc ccgangcccc tgatcagggc 420
cangaagctt gagaagctnc tcggcacgcc ggcgaagat 459

<210> 8
<211> 152
<212> PRT
<213> Zea mays

<220>
<221> VARIANT
<222> (1)...(152)
<223> Xaa = Any Amino Acid

<400> 8

```
Xaa Xaa Xaa Xaa Ala Ala Pro Arg Arg Arg Thr Glu Ala Gly Arg Lys
 1               5                  10                  15
Ala Leu Thr Ser Pro Gly Ala Arg Ser Arg Gln Gln Glu Arg Met Ala
            20                  25                  30
Ala Glu Xaa Ala Ala Ala Ala Thr Leu Arg Thr Ala Leu Ser Gln Ala
        35                  40                  45
Thr Gly Gln Glu Gln Arg Ala Ser Leu Leu Cys Thr Pro Ala Gly Ala
    50                  55                  60
Ser Val Ser Ile Pro Lys Gln Trp Tyr Asn Leu Ile Ala Asp Leu Pro
65                  70                  75                  80
Val Lys Pro Pro Pro Pro Leu His Pro Gln Thr His Gln Pro Leu Asn
                85                  90                  95
Pro Ser Asp Leu Ser Pro Leu Phe Pro Asp Glu Leu Ile Arg Gln Glu
            100                 105                 110
Val Thr Asp Glu Arg Phe Val Asp Ile Pro Glu Glu Val Ile Asp Val
        115                 120                 125
Tyr Lys Leu Trp Arg Pro Xaa Pro Leu Ile Arg Ala Xaa Lys Leu Glu
    130                 135                 140
Lys Leu Leu Gly Thr Pro Ala Lys
145                 150
```

<210> 9
<211> 560
<212> DNA
<213> Zea mays

<220>
<221> misc_feature
<222> (1)...(560)
<223> n = A,T,C or G

<400> 9

cagtccgcag acgagagtgg aagcggggca cagacagaga agccaatggc gctgtccctc 60
tccctcgcgc gccccggccc ctcgccgttt ccgccggcgc aggagccagg aagctacccg 120
cagccagcct cgcattcccg gcgaaatcct tcttcggcgc gccgctggcc gccaccgcgg 180
cctccgtcgc gtcgccgctc ccgcgcaacc ggccacctcc accacctcgc tcgaggtcgt 240
cgcggcgggg aagaagggct acaagatgaa gacgcacaan gctcggcgaa cgttccgggt 300
gacgggaggg gcaagatctg cggcggtgcg cnggaaagca gcattgctcg ccaagaagaa 360
caccaagcgc aagaagaggc tcttcgaaga tggtgcaagt caacaagant gactacgaca 420
atgttacggg tgcaatgccc tacttcaaat tgaataagga aagcaaactg agagctaacg 480
tggtcttcaa aaaatcatta attttcgttc caccaattgt aattttgtgt ancttccact 540

gntatttcct tcttcaaaaa 560

<210> 10
<211> 150
<212> PRT
<213> Zea mays

<220>
<221> VARIANT
<222> (1)...(150)
<223> Xaa = Any Amino Acid

<400> 10

```
Gln Ser Ala Asp Glu Ser Gly Ser Gly Ala Gln Thr Glu Lys Pro Met
 1               5                  10                  15
Ala Leu Ser Leu Ser Leu Ala Arg Pro Gly Pro Ser Pro Phe Pro Pro
            20                  25                  30
Ala Gln Glu Pro Gly Ser Tyr Pro Gln Pro Ala Ser His Ser Arg Arg
        35                  40                  45
Asn Pro Ser Ser Ala Arg Arg Trp Pro Pro Pro Arg Pro Pro Ser Arg
    50                  55                  60
Arg Arg Ser Arg Ala Thr Gly His Leu His His Leu Ala Arg Gly Arg
65                  70                  75                  80
Arg Gly Gly Glu Glu Gly Leu Gln Asp Glu Asp Ala Gln Xaa Ser Ala
                85                  90                  95
Asn Val Pro Gly Asp Gly Arg Gly Lys Ile Cys Gly Gly Ala Xaa Glu
            100                 105                 110
Ser Ser Ile Ala Arg Gln Glu Glu His Gln Ala Gln Glu Glu Ala Leu
        115                 120                 125
Arg Arg Trp Cys Lys Ser Thr Arg Xaa Thr Thr Thr Met Leu Arg Val
    130                 135                 140
Gln Cys Pro Thr Ser Asn
145                 150
```

<210> 11
<211> 461
<212> DNA
<213> Zea mays

<220>
<221> misc_feature
<222> (1)...(461)
<223> n = A,T,C or G

<400> 11

```
gtagcgaacc ngagaggaat ggcgtcgctg tcgtgcgcct ccacgctgcc ctggaccccc   60
gcctcctacc cgcggccctc ggcggggcgc tgcctcacgg cctctcgccg ggctccgtcg  120
ctggtcatcg tcgcccaggg gaaggtcaag aantaccgcc aggncatatt gacagacgac  180
atagaagagg tggggaagaa aggggataca ctgaaggtgc gagctggatt ctaccgcaac  240
ttcctccttc ccaagggcaa ggctacactt ctgaccccag aagtcctcaa ggaaatgcag  300
ctggagcagg agagaataga agctganaag aagcgggtaa aagaagaacg cgcaacaact  360
tgctngagtt tttgaaaacc attggcgnct ttcaaaantc cctcgcaaag gtggaaaagg  420
gaaaacagga tctttgggag tgtnacggca caaggntgtn g                      461
```

<210> 12
<211> 124
<212> PRT

<213> Zea mays

<220>
<221> VARIANT
<222> (1)...(124)
<223> Xaa = Any Amino Acid

<400> 12

```
Val Ala Asn Xaa Arg Gly Met Ala Ser Leu Ser Cys Ala Ser Thr Leu
 1               5                  10                  15
Pro Trp Thr Pro Ala Ser Tyr Pro Arg Pro Ser Ala Gly Arg Cys Leu
            20                  25                  30
Thr Ala Ser Arg Arg Ala Pro Ser Leu Val Ile Val Ala Gln Gly Lys
        35                  40                  45
Val Lys Xaa Tyr Arg Gln Xaa Ile Leu Thr Asp Asp Ile Glu Glu Val
    50                  55                  60
Gly Lys Lys Gly Asp Thr Leu Lys Val Arg Ala Gly Phe Tyr Arg Asn
65                  70                  75                  80
Phe Leu Leu Pro Lys Gly Lys Ala Thr Leu Leu Thr Pro Glu Val Leu
                85                  90                  95
Lys Glu Met Gln Leu Glu Gln Glu Arg Ile Glu Ala Xaa Lys Lys Arg
            100                 105                 110
Val Lys Glu Glu Arg Ala Thr Thr Cys Xaa Ser Phe
        115                 120
```

<210> 13
<211> 437
<212> DNA
<213> Zea mays

<220>
<221> misc_feature
<222> (1)...(437)
<223> n = A,T,C or G

<400> 13

```
nnnnnnnnnn nnnnngagca gctagcagta tacaagagct tgcttgctgt cctagtccta      60
cctactgacg agctcgccat ggcctccctc tcctctgccg ccgtcaccgc cccttccttc     120
gccgcaccag cgccagcccg cgccgtggtc agaaggaggt ccttcaccgt gcgcgcctct     180
ctccgcaagg ccaccggcac cgccgccgtg gcaatggctg ccagcgccct gcttgccggc     240
ggtgccatgg cccaggaggt gctgctgggc gcaggngacg gcgggctcgt cttcgagccc     300
agccagttca ccgtcaaggc cggcgacacc atcacattca agaacaacgc cggcttcccg     360
cacaacgtcg tcttcgacga agaagaggtg ccgaagcggc gtcgaagcca ccaagatctc     420
gcaggaggag tacctta                                                    437
```

<210> 14
<211> 145
<212> PRT
<213> Zea mays

<220>
<221> VARIANT
<222> (1)...(145)
<223> Xaa = Any Amino Acid

<400> 14

```
Xaa Xaa Xaa Xaa Xaa Glu Gln Leu Ala Val Tyr Lys Ser Leu Leu Ala
```

```
 1               5                  10                  15
Val Leu Val Leu Pro Thr Asp Glu Leu Ala Met Ala Ser Leu Ser Ser
            20                  25                  30
Ala Ala Val Thr Ala Pro Ser Phe Ala Ala Pro Ala Pro Ala Arg Ala
        35                  40                  45
Val Val Arg Arg Arg Ser Phe Thr Val Arg Ala Ser Leu Arg Lys Ala
    50                  55                  60
Thr Gly Thr Ala Ala Val Ala Met Ala Ala Ser Ala Leu Leu Ala Gly
65                  70                  75                  80
Gly Ala Met Ala Gln Glu Val Leu Leu Gly Ala Gly Asp Gly Gly Leu
                85                  90                  95
Val Phe Glu Pro Ser Gln Phe Thr Val Lys Ala Gly Asp Thr Ile Thr
            100                 105                 110
Phe Lys Asn Asn Ala Gly Phe Pro His Asn Val Val Phe Asp Glu Glu
        115                 120                 125
Glu Val Pro Lys Arg Arg Arg Ser His Gln Asp Leu Ala Gly Gly Val
    130                 135                 140
Pro
145
```

<210> 15
<211> 463
<212> DNA
<213> Zea mays

<400> 15

```
cggcgtcacc gtcaccgccg ctttccgagc tctcatggcg gcgtccgcgt cctccctcct   60
tgccgccccc gcatcctcat cctgtggggc catctccccc cagctccctc gtggagctcc  120
ggccgctgcc tccgtcgctt cgccgtcgcg tcactcctgc tacctcctcc gcgcttcccc  180
ttcgaggcgc catcggagtc gcttcgtagc caacgccgct cccaccatgc agcctccggc  240
ggagtccagg gtctccacgg tagtcgatgt cgacctcggc gaccgtagct atccgatcta  300
catcggcgca ggcctcctcg acgagccgga cctgctgcag aggcatgttc atggtaagag  360
ggttctggtg gtgaccaaca cgaccgtcgc gccgctttaa ctggacaagg tgacatgggc  420
actcaaccac aacaactgaa tgtatcagtg gaaagcgtga tcc                    463
```

<210> 16
<211> 132
<212> PRT
<213> Zea mays

<400> 16

```
Gly Val Thr Val Thr Ala Ala Phe Arg Ala Leu Met Ala Ala Ser Ala
 1               5                  10                  15
Ser Ser Leu Leu Ala Ala Pro Ala Ser Ser Ser Cys Gly Ala Ile Ser
            20                  25                  30
Pro Gln Leu Pro Arg Gly Ala Pro Ala Ala Ala Ser Val Ala Ser Pro
        35                  40                  45
Ser Arg His Ser Cys Tyr Leu Leu Arg Ala Ser Pro Ser Arg Arg His
    50                  55                  60
Arg Ser Arg Phe Val Ala Asn Ala Ala Pro Thr Met Gln Pro Pro Ala
65                  70                  75                  80
Glu Ser Arg Val Ser Thr Val Val Asp Val Asp Leu Gly Asp Arg Ser
                85                  90                  95
Tyr Pro Ile Tyr Ile Gly Ala Gly Leu Leu Asp Glu Pro Asp Leu Leu
            100                 105                 110
Gln Arg His Val His Gly Lys Arg Val Leu Val Val Thr Asn Thr Thr
        115                 120                 125
```

Val Ala Pro Leu
130

<210> 17
<211> 448
<212> DNA
<213> Zea mays

<220>
<221> misc_feature
<222> (1)...(448)
<223> n = A,T,C or G

<400> 17
gacggcagtc atggcttcca tcaccctcct ctccctcgct cccaccgcca tcttcctcca 60
catcccggct ttcacctcct cctccgtcgt tggtcccgga atcctagccg gccgctgggc 120
cgcgccccgc gccctcccgc tccgcgcgct cccgccccgt cgcgtcaccg tcgtctgcag 180
cagcgccgcg gcggcagccg aggcctcgga cgccgcagcg ccagtggaga agttccggct 240
cgacaatttg gggccccaga aggggtcccg ccggcggccc aagcgtaagg gacgcggtat 300
tgctgcaggc cagggcgcga gctgcgggtt cggtatgcgc gggcagaaat cgcgctcggg 360
gcccggcgtc cggcgcgggt tcganggaag gcagatgccg ctctaacgcc gtctccccaa 420
gtacggggat tncgggcggg atgcacat 448

<210> 18
<211> 134
<212> PRT
<213> Zea mays

<400> 18
Thr Ala Val Met Ala Ser Ile Thr Leu Leu Ser Leu Ala Pro Thr Ala
1 5 10 15
Ile Phe Leu His Ile Pro Ala Phe Thr Ser Ser Ser Val Val Gly Pro
20 25 30
Gly Ile Leu Ala Gly Arg Trp Ala Ala Pro Arg Ala Leu Pro Leu Arg
35 40 45
Ala Leu Pro Pro Arg Arg Val Thr Val Val Cys Ser Ser Ala Ala Ala
50 55 60
Ala Ala Glu Ala Ser Asp Ala Ala Ala Pro Val Glu Lys Phe Arg Leu
65 70 75 80
Asp Asn Leu Gly Pro Gln Lys Gly Ser Arg Arg Arg Pro Lys Arg Lys
85 90 95
Gly Arg Gly Ile Ala Ala Gly Gln Gly Ala Ser Cys Gly Phe Gly Met
100 105 110
Arg Gly Gln Lys Ser Arg Ser Gly Pro Gly Val Arg Arg Gly Phe Xaa
115 120 125
Gly Arg Gln Met Pro Leu
130

<210> 19
<211> 528
<212> DNA
<213> Zea mays

<220>
<221> misc_feature
<222> (1)...(528)
<223> n = A,T,C or G

<400> 19
gccaaacant gcgncacnag nactataacc aaccaggnca accaaacctt ctccgatggc 60
cgcgcccgtg tcgcagccgc cggtgtccgc cagggcgtcc acacggtttc tcccccgcgg 120
gataggcgcg ctcccggagt ccgccccccac gtccctccgg ttatccgtcg gccgccgtcg 180
ccgcgcctcc agcctagagg tgaaggcatc aggaaatgtg ttcgggaact acttccaggt 240
tgcaacctat ggcgaatccc atggaggggg tgttggttgc gttatcagtg gctgcccacc 300
cagaattcct ctcactgagg cagacatgca agtagaactc gatagaagac gtccgggtca 360
aagtagaatt acaacccccaa gaaaggagac tgatacatgc aaaattctat canggacaca 420
tgatgggatg actactggta caccaattca cgtctttgtc ccaaacacag atcaaanggg 480
tggtgattac agtgaaatgt ctaangcgta cagaccatcc catgcaga 528

<210> 20
<211> 166
<212> PRT
<213> Zea mays

<220>
<221> VARIANT
<222> (1)...(166)
<223> Xaa = Any Amino Acid

<400> 20
Pro Thr Arg Xaa Thr Lys Pro Ser Pro Met Ala Ala Pro Val Ser Gln
1 5 10 15
Pro Pro Val Ser Ala Arg Ala Ser Thr Arg Phe Leu Pro Arg Gly Ile
20 25 30
Gly Ala Leu Pro Glu Ser Ala Pro Thr Ser Leu Arg Leu Ser Val Gly
35 40 45
Arg Arg Arg Arg Ala Ser Ser Leu Glu Val Lys Ala Ser Gly Asn Val
50 55 60
Phe Gly Asn Tyr Phe Gln Val Ala Thr Tyr Gly Glu Ser His Gly Gly
65 70 75 80
Gly Val Gly Cys Val Ile Ser Gly Cys Pro Pro Arg Ile Pro Leu Thr
85 90 95
Glu Ala Asp Met Gln Val Glu Leu Asp Arg Arg Arg Pro Gly Gln Ser
100 105 110
Arg Ile Thr Thr Pro Arg Lys Glu Thr Asp Thr Cys Lys Ile Leu Ser
115 120 125
Xaa Thr His Asp Gly Met Thr Thr Gly Thr Pro Ile His Val Phe Val
130 135 140
Pro Asn Thr Asp Gln Xaa Gly Gly Asp Tyr Ser Glu Met Ser Xaa Ala
145 150 155 160
Tyr Arg Pro Ser His Ala
165

<210> 21
<211> 427
<212> DNA
<213> Zea mays

<220>
<221> misc_feature
<222> (1)...(427)
<223> n = A,T,C or G

<400> 21

```
ggacctggcn gagaaggaga cgccggagag cgagctgccc ccgacgttcc tccgcgagga      60
cgggcgcggc gccgggtccg tgcgggcgcg gttcgagtcc atgatacggc gggtgcaggg     120
ggaggtctgc gccgcgctcg aggaggcaga tggcggtggg gcgcgcttcg tcgaggacgt     180
ctggtcgcgc cccggcggcg gcggcggcat cagccgggtc ctgcaagacg gccgcgtctt     240
cgagaaggcc ggggtcaacg tctccgtcgt ctacggggtc atgccaaccg acgcctaacg     300
cgccgccaan ggggaagccg gcaagaacga agccgccgcg gatggccaaa agctggcccc     360
gttgccattc ttcgccgcgg gtattanctt cggtaantng cntttctgct tcaattnaag     420
ttaagaa                                                              427
```

<210> 22
<211> 107
<212> PRT
<213> Zea mays

<220>
<221> VARIANT
<222> (1)...(107)
<223> Xaa = Any Amino Acid

<400> 22

```
Tyr Gly Gly Cys Arg Gly Arg Ser Ala Pro Arg Ser Arg Arg Gln Met
 1               5                  10                  15
Ala Val Gly Arg Ala Ser Ser Arg Thr Ser Gly Arg Ala Pro Ala Ala
            20                  25                  30
Ala Ala Ala Ser Ala Gly Ser Cys Lys Thr Ala Ala Ser Ser Arg Arg
        35                  40                  45
Pro Gly Ser Thr Ser Pro Ser Ser Thr Gly Ser Cys Gln Pro Thr Pro
    50                  55                  60
Asn Ala Pro Pro Xaa Gly Lys Pro Ala Arg Thr Lys Pro Pro Arg Met
65                  70                  75                  80
Ala Lys Ser Trp Pro Arg Cys His Ser Ser Pro Arg Val Leu Xaa Ser
                85                  90                  95
Val Xaa Xaa Xaa Ser Ala Ser Ile Xaa Val Lys
            100                 105
```

<210> 23
<211> 717
<212> DNA
<213> Artificial Sequence

<220>
<221> CDS
<222> (1)...(712)

<223> Green fluorescent protein of jelly fish modified
to have Zea mays preferred codons.

<400> 23

```
atg tcc aag ggc gag gag ctc ttc acc ggc gtg gtg ccc atc ctc gtg       48
Met Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu Val
 1               5                  10                  15

gag ctc gac ggc gac gtg aac ggc cac aag ttc tcc gtg tcc ggc gag       96
Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu
            20                  25                  30

ggc gag ggc gac gcc acc tac ggc aag ctc acc ctc aag ttc atc tgc      144
```

```
Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys
          35                  40                  45

acc acc ggc aag ctc ccc gtg ccc tgg ccc acc ctc gtg acc acc ttc      192
Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Phe
      50                  55                  60

tcc tac ggc gtg cag tgc ttc tcc agg tac ccc gac cac atg aag cag      240
Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln
 65                  70                  75                  80

cac gac ttc ttc aag tca gcc atg ccc gag ggc tac gtg cag gag agg      288
His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg
                 85                  90                  95

acc atc ttc ttc aag gac gac ggc aac tac aag acc agg gcc gag gtg      336
Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val
            100                 105                 110

aag ttc gaa ggc gac acc ctc gtg aac agg att gag ctc aag ggc atc      384
Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile
        115                 120                 125

gac ttc aag gag gac ggc aac atc ctc ggc cac aag ctc gag tac aac      432
Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn
    130                 135                 140

tac aac tcc cac aac gtg tac atc atg gcc gac aag cag aag aac ggc      480
Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly
145                 150                 155                 160

atc aag gtg aac ttc aag atc agg cac aac atc gag gac ggc tca gtg      528
Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val
                165                 170                 175

cag ctc gct gac cac tac cag cag aac acc ccc atc ggc gac ggc ccc      576
Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro
            180                 185                 190

gtg ctc ctc ccc gac aac cac tac ctc tcc acc cag tcc gcc ctc tcc      624
Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser
        195                 200                 205

aag gac ccc aac gag aag agg gac cac atg gtg ctc ctc gag ttc gtg      672
Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val
    210                 215                 220

acc gct gct ggc atc acc cac ggc atg gac gag ctc tac a agtga         717
Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu Leu Tyr
225                 230                 235
```

<210> 24
<211> 237
<212> PRT
<213> Zea mays

<400> 24

```
Met Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu Val
 1               5                  10                  15
Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu
            20                  25                  30
Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys
        35                  40                  45
Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Phe
    50                  55                  60
Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln
65                  70                  75                  80
His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg
                85                  90                  95
Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val
            100                 105                 110
Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile
        115                 120                 125
Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn
    130                 135                 140
Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly
145                 150                 155                 160
Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val
                165                 170                 175
Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro
            180                 185                 190
Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser
        195                 200                 205
Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val
    210                 215                 220
Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu Leu Tyr
225                 230                 235
```

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :
C12N 15/82, C07K 14/415, C12N 5/10, 15/62, A01H 5/00

A3

(11) International Publication Number: WO 00/12732

(43) International Publication Date: 9 March 2000 (09.03.00)

(21) International Application Number: PCT/US99/18955

(22) International Filing Date: 25 August 1999 (25.08.99)

(30) Priority Data:
60/098,225 28 August 1998 (28.08.98) US

(71) Applicant *(for all designated States except US)*: PIONEER HI-BRED INTERNATIONAL, INC. [US/US]; 800 Capital Square, 400 Locust Street, Des Moines, IA 50319 (US).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: BENSEN, Robert, J. [US/US]; 200 Northwest Prairie Creek Drive, Grimes, IA 50111 (US).

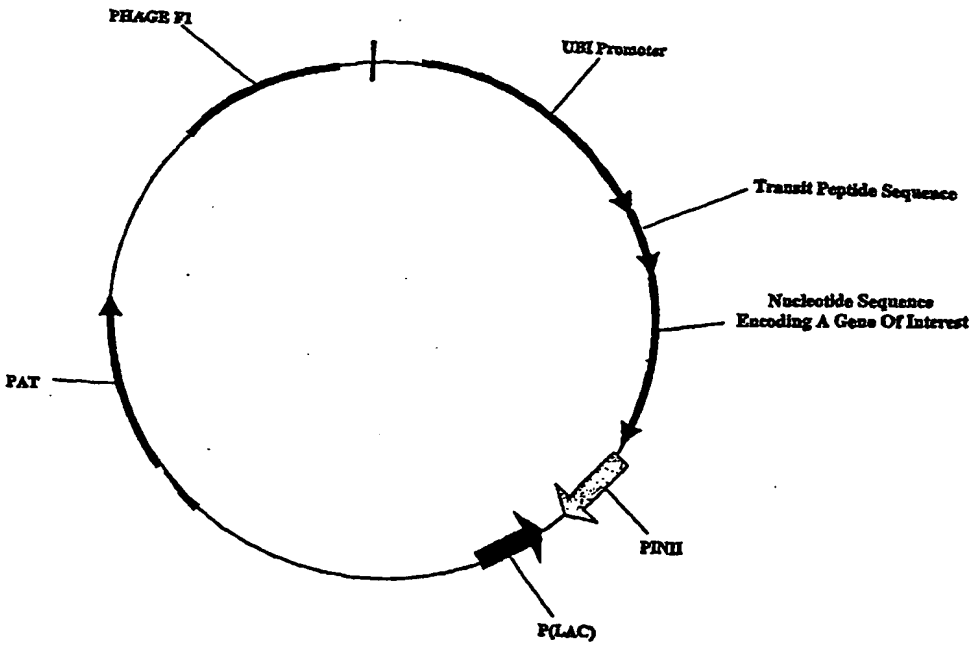
(74) Agents: SPRUILL, W., Murray et al.; Alston & Bird LLP, P.O. Drawer 34009, Charlotte, NC 28234-4009 (US).

(81) Designated States: AE, AL, AM, AT, AT (Utility model), AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, CZ (Utility model), DE, DE (Utility model), DK, DK (Utility model), DM, EE, EE (Utility model), ES, FI, FI (Utility model), GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SK (Utility model), SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(88) Date of publication of the international search report:
19 October 2000 (19.10.00)

(54) Title: ORGANELLE TARGETING SEQUENCES



(57) Abstract

Compositions and methods are provided for modulating the subcellular localization of proteins in a cell. Compositions include nucleotide and amino acid sequences of transit peptide sequences from maize. Such sequences find utility in the enhanced or modified localization of protein to a plastid or compartment thereof.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

INTERNATIONAL SEARCH REPORT

Inter nal Application No
PCT/US 99/18955

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/82 C07K14/415 C12N5/10 C12N15/62 A01H5/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N C07K A01H

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 97 04114 A (RHONE POULENC AGROCHIMIE ;DEROSE RICHARD (FR); CHAUBET NICOLE (FR)) 6 February 1997 (1997-02-06) page 4, line 18 - line 36 page 9, line 7 - line 15 page 15, line 6 -page 16, line 9; claims 15-17,24 ---	1-16
X	DELLA-CIOPPA, G. ET AL.: "Targeting a herbicide-resistant enzyme from Escherichia coli to chloroplasts of higher plants." BIO/TECHNOLOGY, vol. 5, June 1987 (1987-06), pages 579-84, XP002140206 abstract; figure 1A --- -/--	1-16

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
15 June 2000	05. 07. 00

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Smalt, R

INTERNATIONAL SEARCH REPORT

Interr nal Application No
PCT/US 99/18955

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5 639 952 A (QUAIL PETER H ET AL) 17 June 1997 (1997-06-17)	7-16
Y	column 13, line 59 -column 14, line 9	1-6

Y	KAVANAGH T A ET AL: "TARGETING A FOREIGN PROTEIN TO CHLOROPLASTS USING FUSIONS TO THE TRANSIT PEPTIDE OF A CHLOROPHYLL A/B PROTEIN" MOLECULAR AND GENERAL GENETICS,DE,SPRINGER VERLAG, BERLIN, vol. 215, no. 1, 1 December 1988 (1988-12-01), pages 38-45, XP000027953 ISSN: 0026-8925 the whole document	1-6

X	DATABASE EMBL - EMEST2 [Online] Entry/Acc.no. AA661454, 14 November 1997 (1997-11-14) BAYSDORFER C.: "zEST00799 Maize Leaf, Stratagene #937005 Zea mays cDNA clone csuh00799 5' end similar to plastid ribosomal protein CL9, mRNA sequence." XP002140210	7,10, 12-16
Y	the whole document	1-6,8,9, 11

Y	GANTT, J.S.: "Nucleotide sequences of cDNAs encoding four complete nuclear-encoded plastid ribosomal proteins." CURRENT GENETICS, vol. 14, 1988, pages 519-28, XP000914620 figure 3	1-6,8,9, 11

X	DATABASE NCBI - DBEST [Online] Acc.no. 1738932, 5 June 1998 (1998-06-05) WEN, T.J. ET AL.: "MEST6-D3.TW1412.Seq ISUM2 Zea mays cDNA clone MEST6-D3 5', mRNA sequence." retrieved from HTTP://WWW.NCBI.NLM.NIH.GOV/IRX/CGI-BIN/BIRX_DOC?DBEST+1641583 XP002140211	7,10, 12-16
Y	the whole document	1-6,8,9, 11

Y	SMOOKER, P.M. ET AL.: "Ribosomal protein L35: identification in spinach chloroplasts and isolation of a cDNA clone encoding its cytoplasmic precursor." BIOCHEMISTRY, vol. 29, 1990, pages 9733-6, XP002140207 the whole document	1-6,8,9, 11

	-/--	

INTERNATIONAL SEARCH REPORT

Interr ıal Application No
PCT/US 99/18955

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE NCBI - DBEST [Online] Acc.no. 1716328, 26 May 1998 (1998-05-26) WEN, T.J. ET AL.: "MEST2-D1.TW1412.Seq ISUM2 Zea mays cDNA clone MEST2-D1 5', mRNA sequence." retrieved from HTTP://WWW.NCBI.NLM.NIH.GOV/IRX/CGI-BIN/BIRX_DOC?DBEST+1621781 XP002140212	7,10, 12-16
Y	the whole document	1-6,8,9, 11

Y	SMEEKENS, S. ET AL.: "Import into chloroplasts of a yeast mitochondrial protein directed by ferrodoxin and plastocyanin transit peptides." PLANT MOLECULAR BIOLOGY, vol. 9, 1987, pages 377-88, XP002140208 the whole document	1-6,8,9, 11

X	DATABASE NCBI - DBEST [Online] Id: 1290373, Acc.no. C72774, 22 September 1997 (1997-09-22) SASAKI, T.: "Rice cDNA, partial sequence (E2219_1A)." retrieved from HTTP://WWW.NCBI.NLM.NIH.GOV/IRX/CGI-BIN/BIRX_DOC?DBEST+1216681 XP002140213	7,10, 12-16
Y	the whole document	1-6,8,9, 11

Y	GÖRLACH, J. ET AL.: "Differential expression of tomato (Lycopersicon esculentum L.) genes encoding shikimate pathway isoenzymes. II. Chorismate synthase." PLANT MOLECULAR BIOLOGY, vol. 23, 1993, pages 707-16, XP002140209 abstract page 771, left-hand column, line 9 - line 25	1-6,8,9, 11

X	KRUSE, E. ET AL.: "Coproporphyrinogen III oxidase from barley and tobacco - sequence analysis and initial expression studies." PLANTA, vol. 196, 1995, pages 796-803, XP000920807	7,10, 12-16
Y	page 798, left-hand column, paragraph 4 -right-hand column, paragraph 4; figure 1	1-6,8,9, 11

	-/--	

INTERNATIONAL SEARCH REPORT

Inter nal Application No
PCT/US 99/18955

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 98 20144 A (HAWKES TIMOTHY ROBERT ;KNIGHT MARY ELIZABETH (GB); THOMPSON PAUL A) 14 May 1998 (1998-05-14) the whole document, particularly page 28, lines 1-24	1-6,8,9, 11
A	EP 0 508 909 A (RHONE POULENC AGROCHIMIE) 14 October 1992 (1992-10-14) the whole document	
A	EP 0 337 899 A (RHONE POULENC AGROCHIMIE) 18 October 1989 (1989-10-18) the whole document	
A	WO 91 03561 A (DU PONT) 21 March 1991 (1991-03-21) claims 20,21	
A	VON HEIJNE, G. ET AL.: "Domain structure of mitochondrial and chloroplast targeting peptides" EUROPEAN JOURNAL OF BIOCHEMISTRY, vol. 180, 1989, pages 535-45, XP000877203 cited in the application the whole document	
A	ARCHER, E.K. ET AL.: "Current views on chloroplast protein import and hypotheses on the origin of the transport mechanism." JOURNAL OF BIOENGERGETICS AND BIOMEMBRANES, vol. 22, 1990, pages 789-810, XP000877205 cited in the application the whole document	
A	LEROUX, B. ET AL.: "Engineering herbicide resistance in tobacco plants by expression of a bromoxynil specific nitrilase." BULLETIN DE LA SOCIETE BOTANIQUE DE FRANCE, vol. 137, no. 3/4, 1990, pages 65-78, XP000884652 the whole document	
E	WO 00 05387 A (KINNEY ANTHONY JOHN ;RAFALSKI J ANTONI (US); DU PONT (US); RENDINA) 3 February 2000 (2000-02-03) seq.ID.1 and 2	7-16
E	WO 00 05353 A (DU PONT ;CAHOON REBECCA E (US); FALCO SAVERIO CARL (US); PEMBER ST) 3 February 2000 (2000-02-03) seq.ID.1 and 2	7-16

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 99/18955

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheets

1. ☒ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☒ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-16, al partially

Method for modulating cellular localization of a protein through the use of a transit peptide with the amino acid sequence of seq.ID.2, or encoded by seq.ID.1 or a nucleic acid which hybridizes thereto under stringent conditions, such an isolated nucleic acid, expression cassette comprising it, vector comprising the cassette, transformed plant comprising the vector, seed of said plant, and isolated protein comprising said transit peptide or a homologue thereof encoded by said nucleic acids.

2. Claims: 1-16, al partially

Method for modulating cellular localization of a protein through the use of a transit peptide with the amino acid sequence of seq.ID.4, or encoded by seq.ID.3 or a nucleic acid which hybridizes thereto under stringent conditions, such an isolated nucleic acid, expression cassette comprising it, vector comprising the cassette, transformed plant comprising the vector, seed of said plant, and isolated protein comprising said transit peptide or a homologue thereof encoded by said nucleic acids.

3. Claims: 1-16, al partially

Method for modulating cellular localization of a protein through the use of a transit peptide with the amino acid sequence of seq.ID.6, or encoded by seq.ID.5 or a nucleic acid which hybridizes thereto under stringent conditions, such an isolated nucleic acid, expression cassette comprising it, vector comprising the cassette, transformed plant comprising the vector, seed of said plant, and isolated protein comprising said transit peptide or a homologue thereof encoded by said nucleic acids.

4. Claims: 1-16, al partially

Method for modulating cellular localization of a protein through the use of a transit peptide with the amino acid sequence of seq.ID.8, or encoded by seq.ID.7 or a nucleic acid which hybridizes thereto under stringent conditions, such an isolated nucleic acid, expression cassette comprising it, vector comprising the cassette, transformed plant comprising the vector, seed of said plant, and isolated protein comprising said transit peptide or a homologue thereof encoded by said nucleic acids.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

5. Claims: 1-16, al partially

Method for modulating cellular localization of a protein through the use of a transit peptide with the amino acid sequence of seq.ID.10, or encoded by seq.ID.9 or a nucleic acid which hybridizes thereto under stringent conditions, such an isolated nucleic acid, expression cassette comprising it, vector comprising the cassette, transformed plant comprising the vector, seed of said plant, and isolated protein comprising said transit peptide or a homologue thereof encoded by said nucleic acids.

6. Claims: 1-16, al partially

Method for modulating cellular localization of a protein through the use of a transit peptide with the amino acid sequence of seq.ID.12, or encoded by seq.ID.11 or a nucleic acid which hybridizes thereto under stringent conditions, such an isolated nucleic acid, expression cassette comprising it, vector comprising the cassette, transformed plant comprising the vector, seed of said plant, and isolated protein comprising said transit peptide or a homologue thereof encoded by said nucleic acids.

7. Claims: 1-16, al partially

Method for modulating cellular localization of a protein through the use of a transit peptide with the amino acid sequence of seq.ID.14, or encoded by seq.ID.13 or a nucleic acid which hybridizes thereto under stringent conditions, such an isolated nucleic acid, expression cassette comprising it, vector comprising the cassette, transformed plant comprising the vector, seed of said plant, and isolated protein comprising said transit peptide or a homologue thereof encoded by said nucleic acids.

8. Claims: 1-16, al partially

Method for modulating cellular localization of a protein through the use of a transit peptide with the amino acid sequence of seq.ID.16, or encoded by seq.ID.15 or a nucleic acid which hybridizes thereto under stringent conditions, such an isolated nucleic acid, expression cassette comprising it, vector comprising the cassette, transformed plant comprising the vector, seed of said plant, and isolated protein comprising said transit peptide or a homologue thereof encoded by said nucleic acids.

9. Claims: 1-16, al partially

Method for modulating cellular localization of a protein

through the use of a transit peptide with the amino acid sequence of seq.ID.18, or encoded by seq.ID.17 or a nucleic acid which hybridizes thereto under stringent conditions, such an isolated nucleic acid, expression cassette comprising it, vector comprising the cassette, transformed plant comprising the vector, seed of said plant, and isolated protein comprising said transit peptide or a homologue thereof encoded by said nucleic acids.

10. Claims: 1-16, al partially

Method for modulating cellular localization of a protein through the use of a transit peptide with the amino acid sequence of seq.ID.20, or encoded by seq.ID.19 or a nucleic acid which hybridizes thereto under stringent conditions, such an isolated nucleic acid, expression cassette comprising it, vector comprising the cassette, transformed plant comprising the vector, seed of said plant, and isolated protein comprising said transit peptide or a homologue thereof encoded by said nucleic acids.

11. Claims: 1-16, al partially

Method for modulating cellular localization of a protein through the use of a transit peptide with the amino acid sequence of seq.ID.22, or encoded by seq.ID.21 or a nucleic acid which hybridizes thereto under stringent conditions, such an isolated nucleic acid, expression cassette comprising it, vector comprising the cassette, transformed plant comprising the vector, seed of said plant, and isolated protein comprising said transit peptide or a homologue thereof encoded by said nucleic acids.

INTERNATIONAL SEARCH REPORT

...rmation on patent family members

Interr ...nal Application No
PCT/US 99/18955

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9704114 A	06-02-1997	FR 2736929 A	24-01-1997
		AU 718521 B	13-04-2000
		AU 6618496 A	18-02-1997
		BR 9609774 A	26-01-1999
		CA 2223881 A	06-02-1997
		EP 0850311 A	01-07-1998
		JP 11509414 T	24-08-1999
US 5639952 A	17-06-1997	US 5656496 A	12-08-1997
WO 9820144 A	14-05-1998	AU 4789597 A	29-05-1998
		BG 103462 A	31-01-2000
		BR 9712695 A	19-10-1999
		CN 1236394 A	24-11-1999
		CZ 9901616 A	11-08-1999
		EP 0946737 A	06-10-1999
		PL 333131 A	22-11-1999
		SK 60999 A	08-10-1999
EP 0508909 A	14-10-1992	FR 2673643 A	11-09-1992
		AT 169338 T	15-08-1998
		AU 652610 B	01-09-1994
		AU 1144292 A	10-09-1992
		BR 9200790 A	17-11-1992
		CA 2061636 A	06-09-1992
		DE 69226466 D	10-09-1998
		DE 69226466 T	01-04-1999
		EP 0924299 A	23-06-1999
		ES 2118802 T	01-10-1998
		IL 101115 A	10-01-1997
		JP 5095789 A	20-04-1993
		MX 9200915 A	01-09-1992
		US RE36449 E	14-12-1999
		US 5510471 A	23-04-1996
		US 5633448 A	27-05-1997
		ZA 9201645 A	25-11-1992
EP 0337899 A	18-10-1989	FR 2629098 A	29-09-1989
		AT 188250 T	15-01-2000
		AU 629539 B	08-10-1992
		AU 3159189 A	28-09-1989
		BR 8901499 A	14-11-1989
		CA 1314505 A	16-03-1993
		CN 1036038 A,B	04-10-1989
		DD 294040 A	19-09-1991
		DE 68929125 D	03-02-2000
		DE 68929125 T	31-05-2000
		DK 144689 A	24-09-1989
		EG 20026 A	30-10-1997
		EP 0894865 A	03-02-1999
		ES 2139565 T	16-02-2000
		HU 52158 A,B	28-06-1990
		IL 89614 A	19-01-1996
		JP 2009374 A	12-01-1990
		KR 145703 B	01-08-1998
		NZ 228458 A	25-06-1991
		PL 278448 A	27-12-1989
		RO 107001 A	30-08-1993

INTERNATIONAL SEARCH REPORT

mation on patent family members

Interr nal Application No
PCT/US 99/18955

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0337899 A		TR 23682 A	06-06-1990
		US 5559024 A	24-09-1996
		US 5773698 A	30-06-1998
		ZA 8902146 A	29-11-1989
WO 9103561 A	21-03-1991	US 5212296 A	18-05-1993
		AT 133201 T	15-02-1996
		AU 648036 B	14-04-1994
		AU 6272990 A	08-04-1991
		BR 9007656 A	09-03-1993
		CA 2065439 A	12-03-1991
		DE 69024979 D	29-02-1996
		DE 69024979 T	17-10-1996
		DK 554240 T	03-06-1996
		EP 0554240 A	11-08-1993
		ES 2082862 T	01-04-1996
		JP 5500002 T	14-01-1993
		US 5349127 A	20-09-1994
WO 0005387 A	03-02-2000	AU 5113899 A	14-02-2000
		AU 5113999 A	14-02-2000
		AU 5218799 A	14-02-2000
		WO 0005386 A	03-02-2000
		WO 0005353 A	03-02-2000
WO 0005353 A	03-02-2000	AU 5113899 A	14-02-2000
		AU 5113999 A	14-02-2000
		AU 5218799 A	14-02-2000
		WO 0005386 A	03-02-2000
		WO 0005387 A	03-02-2000